VISIT…

LANZAROTE
Caliente.com

ELSEVIER

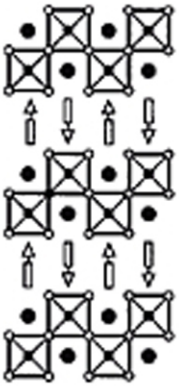
CONCISE ENCYCLOPEDIA OF

COMPOSITE MATERIALS

SECOND EDITION

Editor

A. MORTENSEN



CONCISE ENCYCLOPEDIA OF

COMPOSITE MATERIALS

SECOND EDITION

This Page Intentionally Left Blank

CONCISE ENCYCLOPEDIA OF
COMPOSITE MATERIALS

SECOND EDITION

Editor

A. MORTENSEN

Ecole Polytechnique Fédérale de Lausanne,
Lausanne, Switzerland

ELSEVIER

Amsterdam—Boston—Heidelberg—London—New York—Oxford
Paris—San Diego—San Francisco—Singapore—Sydney—Tokyo

Elsevier
The Boulevard, Langford Lane, Kidlington, Oxford OX5 1GB, UK
Radarweg 29, PO Box 211, 1000 AE Amsterdam, The Netherlands

Second edition 2007

Copyright © 2001–2007 Elsevier Ltd. All rights reserved

No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means electronic, mechanical, photocopying, recording or otherwise without the prior written permission of the publisher

Permissions may be sought directly from Elsevier's Science & Technology Rights Department in Oxford, UK: phone (+44) (0) 1865 843830; fax (+44) (0) 1865 853333; email: permissions@elsevier.com. Alternatively you can submit your request online by visiting the Elsevier web site at http://elsevier.com/locate/permissions, and selecting *Obtaining permission to use Elsevier material*

Notice
No responsibility is assumed by the publisher for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions or ideas contained in the material herein. Because of rapid advances in the medical sciences, in particular, independent verification of diagnoses and drug dosages should be made

ISBN-13: 978-0-08-045126-8
ISBN-10: 0-08-045126-8

For information on all Elsevier publications
visit our website at books.elsevier.com

Printed and bound in The Netherlands

07 08 09 10 11 10 9 8 7 6 5 4 3 2 1

Working together to grow
libraries in developing countries
www.elsevier.com | www.bookaid.org | www.sabre.org
ELSEVIER BOOK AID International Sabre Foundation

CONTENTS

Editor's Preface vii

A Note on Using the Encyclopedia ix

Alphabetical List of Articles xi

Topical List of Articles xiii

Introduction – Composite Materials: Overview xvii

Articles 1

List of Contributors 947

Subject Index 959

DEDICATION

This book is dedicated to the memory of Professor Bernhard Ilschner
of the Ecole Polytechnique Fédérale de Lausanne,
who was among the six Editors-in-Chief
of the Encyclopedia of Materials: Science and Technology.

EDITOR'S PREFACE

This book is a "remake" of a successful predecessor: it follows, in nature, format and ambition, the *Concise Encyclopedia of Composite Materials* that was edited in 1989, and then re-edited in 1994, by Anthony Kelly using the 1986 *Encyclopedia of Material Science and Engineering* as a starting point.

Like the first *Concise Encyclopedia of Composite Materials*, the present volume is a collection of articles from a much larger Encyclopedia. Entries presented here come from the 2001 *Encyclopedia of Materials: Science and Technology*, an entirely rewritten successor to the 1986 *Encyclopedia of Materials Science and Engineering*. The wider scope of the new Encyclopedia of Materials, together with the simple fact that it is more up-to-date, motivated this new edition of a smaller, more easily distributed and shelved, "concise" encyclopedia focused on composite materials.

Composite materials have become so ubiquitous and important a class of materials that many books now exist on the subject. Some offer wide and advanced level coverage of the subject; *Comprehensive Composite Materials*, an extensive six-volume treatise edited by A. Kelly and C. Zweben, which covers in-depth nearly all facets of these variegated materials, stands out in this category. Other books provide more concise and introductory coverage of the subject, while some focus on specific subtopics such as the mechanics of composites.

This book is different; it stands out in two respects.

The first is that it enables quick reference. By virtue of its format, this volume is very easy to search. Many subjects, related to composites but quite disparate, are presented in an independent and brief fashion, contrary to what is provided by a monograph where subject matter is presented incrementally, or a treatise where subject matter is presented exhaustively. The reader interested in rapidly learning essentials of, say, rubber tires (one of the most impressive structural composites if one thinks of their service conditions) or the plastic deformation of cellular materials, will find this book very useful. This, of course, is the main purpose of any encyclopedia.

The second distinguishing feature of this book is more specific to its origin. The fact that it was constructed by gleaning articles from a much wider collection makes it possible to define broadly the subject and its coverage. This Concise Encyclopedia thus contains many bridges between its main subject, composite materials, and the wider discipline of Materials Science and Engineering. Choosing articles from an Encyclopedia containing around two thousand entries makes it possible to give access to topics that are clearly relevant, but often considered peripheral, to composites. One example of this is in the rapidly growing discipline of Biomaterials. Nature's materials are often composites, in both their making and their structure. There are many entries on Biomaterials in the *Encyclopedia of Materials: Science and Technology*; a selection of these, relevant to composites, is included herein. Another example is given by a property such as creep: coverage includes dispersion and particle hardening in a single article; this provides a bridge to the broader subject of metal creep. A third is in applications: some important but less conventional composites will be found here, such as metal-reinforced polymers (used as electrically conductive adhesives) or liquid-crystal containing polymers (used in visual displays). A fourth example is given by highly porous materials: these are often engineered composites of a solid and a gas; several articles in this book concern these materials.

The scope and size of the *Encyclopedia of Materials: Science and Technology* are so large that it has not been possible to fit into this single volume all entries on the subject matter of this book while keeping its size reasonable. Decisions therefore had to made in compiling and editing the selection of articles included herein. The reader should thus be forewarned that these decisions were subjective and that there are other articles relevant to composites as defined here within the *Encyclopedia of Materials: Science and Technology*.

The unique breadth in coverage of composite materials that the reader will find in this single volume was made possible by pooling the contributions of many different authors and editors. Articles concerning the more classical "core" facets of composite materials were mostly commissioned and edited by the undersigned; however, many other articles that the reader will find here were defined and edited by his colleagues among the sixty-two "Subject Editors" of the *Encyclopedia of Materials: Science and Technology*. Their important contribution is herewith acknowledged, all the while pointing out that only the undersigned bears responsibility for the (obviously debatable) choice of articles included herein.

Andreas Mortensen
Lausanne, 2006

This Page Intentionally Left Blank

A NOTE ON USING THE ENCYCLOPEDIA

Like its predecessor, this book begins with an introductory article on composite materials, entitled *Composite Materials: Overview* (see page xvii). All other articles are presented alphabetically by title and an alphabetic listing of the titles is given on page xi.

The title of an article provides a first guide towards researching a given topic. The multilevel Subject Index at the end of the book offers more detailed guidance. An alphabetic list of the contributors names and their affiliations, is also provided as part of the Endmatter in the book.

As in the full *Encyclopedia of Materials: Science and Technology* (EMSAT), articles contain cross references to related articles. Cross references occur both within the text, and in some cases as a listing near the end of an article (in a "See also" section).

Most cross references refer to other articles within the present volume – these are marked by an asterisk (*). Cross references that refer to articles held only in the parent collection (in EMSAT, available for example through Science Direct) are marked by a hash sign (#). All articles end with a Bibliography that gives other sources for further information.

Since Composite Materials is a diverse subject and covers many classes of materials, another point of access to the range of topics covered within a particular subject area is provided by the Topical Listing of Articles (starting on page xiii). Here titles are included in a two-level subject "map" of the contents of this volume.

This Page Intentionally Left Blank

ALPHABETICAL LIST OF ARTICLES

Aircraft Materials
Alternative Techniques for Manufacturing Composite Aerospace Designs, Cost of
Asbestos
Automotive Body Materials
Automotive Chassis/Suspension Materials
Bearing Materials: Plain Bearings
Biological Materials: Synthesis, Structure, and Properties
Biomimetic Materials: Properties and Processing
Bitumens: Modified
Block Copolymers as Precursors to Porous Materials
Block Copolymers as Templates for Functional Materials
Bone and Natural Composites: Properties
Bone Augmentation and Repair
Buildings: Plastics and Composites
Carbon Aerogels
Carbon Blacks
Carbon Fibers
Carbon Nanofibers
Carbon Nanotubes
Carbon–Carbon Composites
Ceramic Fibers from Polymer Precursors
Ceramic Matrix Composites with Roughly Equiaxed Reinforcements: Microstructure and Mechanical Behavior
Ceramic Matrix Composites: Applications
Ceramic Matrix Composites: Matrices and Processing
Ceramic Nanocomposites with Organic Phases, Optics of
Ceramic-modified High-temperature Polymers
Ceramics and Glasses, Sol–Gel Synthesis of
Ceramics: Whisker Toughening
Cermets and Hardmetals
Clay-based Polymer Nanocomposites
Composite Dental Materials: Wear
Composite Materials, Microstructural Design of
Composite Materials: Environmental Effects
Composites for Biomedical Applications
Composites for Sensors and Actuators
Composites, Joining of
Composites, Microstructure of: Quantitative Description
Composites, Physical Properties of
Composites: Interfaces
Concrete as a Building Material
Concrete: Failure Mechanics
Construction: Cellular Materials
Construction Materials: Lightweight Aggregates
Continuous Parallel Fiber Composites: Deformation and Strength
Continuous Parallel Fiber Composites: Fracture
Conventional and Super Abrasive Materials
Creep and Creep-fatigue of Metal-Matrix Composites
Creep of Bituminous Aggregates
Creep of Concrete
Creep of Particle Reinforced Materials
CVD Monofilaments
Designing with Composites
Elastic Behavior of Cellular Solids
Elastic Behavior of Composites
Elastic Structures in Design
Elasticity in Wood and Concrete: Hygromechanical Effects
Elastomers, Ceramic-modified
Elastomers, Reinforcement of
Electronic Packaging: Conductive Adhesives
Electronic Packaging: Elastomer Conductive Polymers
Electronic Packaging: Heat Sink Materials
Fatigue of Particle Reinforced Materials
Fiber Metal Laminates, Fatigue of
Fiber Orientation in Composites: Influence of Fluid Flow
Fiber Strength
Fiberglass
Fibers: Superabsorbent
Fibrous Reinforcements for Composites: Overview
Foams, Microrheology of
Functionally Graded Materials
Glass Fibers
Hard Tissues, Mechanical Properties of
High-performance Fibers
Inorganic and Inorganic–Organic Aerogels
Inorganic Nanotube Materials
Laminates: Physical and Mechanical Behavior
Liquid Crystalline Polymers, Dispersed
Lumber: Laminated Veneer
Magnets: Bonded Permanent Magnets
Materials Selection and Mechanical Design
Mechanical Alloying
Mechanical Testing Methods of Fibers and Composites
Membrane-based Polymer–Ceramic Nanocomposites
Mesoporous Molecular Sieves
Metal Matrix Composites with Roughly Equiaxed Reinforcements: Microstructure and Mechanical Behavior
Metal Matrix Composites, Recycling of

Metal Matrix Composites: Applications
Metal Matrix Composites: Matrices and Processing
Metallic Filaments
Mineral-filled Polymers
Nanoscale Ceramic Composites
Natural Cellulose Fibers and Membranes: Biosynthesis
Natural Cellulose Fibers: Properties
Natural Protein Fibers
Optical Properties of Functional Hybrid Organic–Inorganic Nanocomposites
Organic Aerogels
Organic Analogues to Zeolites and Mesoporous Sieves
Organic–Inorganic Composite Crystals
Paper Products: Classification
Paper: Creep
Paper: Effects of Moisture and Temperature
Paper: Pulping and Bleaching
Paper: Strength and Stiffness
Paper: Structure
Papermaking
PES and PEEK
Plastic Deformation of Cellular Materials
Polymer Composites, Friction and Wear of
Polymer Fiber Processing: High-performance Fibers
Polymer Fibers: Formation and Structure
Polymer Matrix Composites with Roughly Equiaxed Reinforcements: Microstructure and Mechanical Behavior
Polymer Matrix Composites, Recycling of
Polymer Matrix Composites: Applications
Polymer Matrix Composites: Matrices and Processing
Polymer Melt Mixing: Agglomerate Dispersion
Polymer Modified Wood
Polymer Nanocomposites with Metal Dispersions
Polymer Transcrystallinity in Composites
Polymer-based Nanocomposites by Sol–Gel Routes, Applications of
Polymer–Ceramic Nanocomposites: Catalysts
Polymer–Ceramic Nanocomposites: Ceramic Phases
Polymer–Ceramic Nanocomposites: Control of Shrinkage
Polymer–Ceramic Nanocomposites: Interfacial Bonding Agents
Polymer–Ceramic Nanocomposites: Polymer Overview
Polymer-modified Ceramics
Polymer–Nonsilica Ceramic Nanocomposites
Polymer–Silica Nanocomposites
Polymers in Constrained Ceramic Environments
Porous Ceramic Processing
Porous Silica Xerogel
Portland Cements
Powders, Solution Synthesis of
Processing Flexible Polymers to High Performance Fibers
Processing Rigid Polymers to High Performance Fibers
Pulp and Paper: Nonfibrous Components
Radiation Effects in Carbon–Carbon Composites
Reaction Forming
Reactive Ceramic Nanocomposites with Organic and Bio-organic Phases
Reinforced Concrete
Roofing Materials
Rubber Tires
Sporting Materials: Ski Equipment
Spun (Slurry and Sol–Gel) Ceramic Fibers
Superconducting Wires and Cables: Materials and Processing
Supported Catalysts
Textile Fibers: A Comparative Overview
Textile Fibers: Mechanical Properties
Thermally Contracting Materials
Whiskers
Wood Composites: Mineral-bonded
Wood Products: Decay During Use
Wood Products: Weathering
Wood, Constituents of
Wood: Creep and Creep Rupture
Wood: Macroscopic Anatomy
Wood: Moisture Content, Hygroscopicity, and Sorption
Wood: Nonstructural Panel Processes
Wood: Nonstructural Panels
Wood: Strength and Stiffness
Wood: Structural Panel Processes
Wood: Structural Panels
Wood: Thermal Properties
Wood: Ultrastructure
Wood–Plastic Composites
Xerogels
Zeolites

TOPICAL LIST OF ARTICLES

I – INTRODUCTION

a) What is a composite?

Composite Materials: Overview (in preliminary pages)

b) Designing composite materials

Composite Materials, Microsturctural Design of
Composites for Sensors and Actuators
Designing with Composites
Elastic Structures in Design
Electronic Packaging: Heat Sink Materials
Functionally Graded Materials
Materials Selection and Mechanical Design

II – REINFORCEMENTS

a) Fibers and whiskers

Asbestos
Carbon Fibers
Carbon Nanofibers
Carbon Nanotubes
Ceramic Fibers from Polymer Precursors
Ceramics and Glasses, Sol-Gel Synthesis of
CVD Monofilaments
Fiber Strength
Fibrous Reinforcements for Composites: Overview
Glass Fibers
High Performance Fibers
Inorganic Nanotube Materials
Metallic Filaments
Natural Protein Fibers
Polymer Fiber Processing: High-performance Fibers
Polymer Fibers: Formation and Structure
Processing Flexible Polymers to High Performance Fibers
Processing Rigid Polymers to High Performance Fibers
Spun (Slurry and Sol-Gel) Ceramic Fibers
Textile Fibers: A Comparative Overview
Whiskers

b) Particulate

Ceramics and Glasses, Sol-Gel Synthesis of
Conventional and Super Abrasive Materials
Mineral-filled Polymers
Powders, Solution Synthesis of
Thermally Contracting Materials

III – MATRICES AND PROCESSING

Alternative Techniques for Manufacturing Composite Aerospace Designs, Cost of
Ceramic Matrix Composites: Matrices and Processing
Ceramic-modified High-temperature Polymers,
Ceramics and Glasses, Sol-Gel Synthesis of
Clay-based Polymer Nanocomposites
Composites, Joining of
Elastomers, Ceramic-modified
Fiber Orientation in Composites: Influence of Fluid Flow
Mechanical Alloying
Metal Matrix Composites, Recycling of
Metal Matrix Composites: Matrices and Processing
Polymer Matrix Composites, Recycling of
Polymer Melt Mixing: Agglomerate Dispersion
Polymer-Ceramic Nanocomposites: Ceramic Phases
Polymer-Ceramic Nanocomposites: Control of Shrinkage
Reaction Forming
Reactive Ceramic Nanocomposites with Organic and Bio-organic Phases

IV – COMPOSITES BY CLASS

a) Polymer matrix composites

Composites for Sensors and Actuators
Elastomers, Reinforcement of
Electronic Packaging: Conductive Adhesives
Electronic Packaging: Elastomer Conductive Polymers
Fiber Metal Laminates, Fatigue of
Liquid Crystalline Polymers, Dispersed
Magnets: Bonded Permanent Magnets
Mineral-filled Polymers
PES and PEEK
Polymer Composites, Friction and Wear of
Polymer Matrix Composites with Roughly Equiaxed Reinforcements: Microstructure and Mechanical Behavior
Polymer Matrix Composites, Recycling of
Polymer Matrix Composites: Applications
Polymer Transcrystallinity in Composites
Rubber Tires

b) Metal matrix composites

Automotive Chassis/Suspension Materials
Cermets and Hardmetals
Creep and Creep-fatigue of Metal-Matrix Composites
Creep of Particle Reinforced Materials
Electronic Packaging: Heat Sink Materials
Fatigue of Particle Reinforced Materials
Metal Matrix Composites with Roughly Equiaxed Reinforcements: Microstructure and Mechanical Behavior
Metal Matrix Composites, Recycling of
Metal Matrix Composites: Applications
Metal Matrix Composites: Matrices and Processing

Superconducting Wires and Cables: Materials and Processing

c) Ceramic matrix composites

Ceramic Matrix Composites with Roughly Equiaxed Reinforcements: Microstructure and Mechanical Behavior
Ceramic Matrix Composites: Applications
Ceramic Matrix Composites: Matrices and Processing
Ceramics: Whisker Toughening
Nanoscale Ceramic Composites

d) Carbon matrix composites

Carbon-Carbon Composites
Radiation Effects in Carbon-Carbon Composites

e) Biocomposites

Biological Materials: Synthesis, Structure, and Properties
Biomimetic Materials: Properties and Processing
Bone and Natural Composites: Properties
Bone Augmentation and Repair
Composite Dental Materials: Wear
Composites for Biomedical Applications
PES and PEEK

f) Nanocomposites

Carbon Nanofibers
Carbon Nanotubes
Ceramic Nanocomposites with Organic Phases, Optics of
Ceramic-modified High-temperature Polymers
Clay-based Polymer Nanocomposites
Elastomers, Ceramic-modified
Inorganic Nanotube Materials
Mechanical Alloying
Membrane-based Polymer-Ceramic Nanocomposites
Nanoscale Ceramic Composites
Optical Properties of Functional Hybrid Organic-Inorganic Nanocomposites
Polymer Nanocomposites with Metal Dispersions
Polymer-based Nanocomposites by Sol-Gel Routes, Applications of
Polymer-Ceramic Nanocomposites: Catalysts
Polymer-Ceramic Nanocomposites: Ceramic Phases
Polymer-Ceramic Nanocomposites: Control of Shrinkage
Polymer-Ceramic Nanocomposites: Interfacial Bonding Agents
Polymer-Ceramic Nanocomposites: Polymer Overview
Polymer-modified Ceramics
Polymer-Nonsilica Ceramic Nanocomposites
Polymers in Constrained Ceramic Environments
Polymer-Silica Nanocomposites
Reactive Ceramic Nanocomposites with Organic and Bio-organic Phases

g) Wood composites

Lumber: Laminated Veneer
Wood Composites: Mineral-bonded
Wood: Nonstructural Panel Processes
Wood: Nonstructural Panels
Wood: Structural Panel Processes
Wood: Structural Panels
Wood-Plastic Composites

h) Cement composites

Concrete as a Building Material
Concrete: Failure Mechanics
Creep of Concrete
Portland Cements
Reinforced Concrete
Wood Composites: Mineral-bonded

i) Microcellular materials

Bone and Natural Composites: Properties
Bone Augmentation and Repair
Construction: Cellular Materials
Elastic Behavior of Cellular Solids
Materials Selection and Mechanical Design
Plastic Deformation of Cellular Materials
Porous Ceramic Processing

V – PROPERTIES AND CHARACTERIZATION

a) Physical properties

Ceramic Nanocomposites with Organic Phases, Optics of
Composite Materials: Environmental Effects
Composites, Physical Properties of
Laminates: Physical and Mechanical Behavior
Liquid Crystalline Polymers, Dispersed
Magnets: Bonded Permanent Magnets
Superconducting Wires and Cables: Materials and Processing

b) Mechanical and environmental behaviour

Ceramic Matrix Composites with Roughly Equiaxed Reinforcements: Microstructure and Mechanical Behavior
Ceramics: Whisker Toughening
Composite Dental Materials: Wear
Composite Materials: Environmental Effects
Composites: Interfaces
Concrete: Failure Mechanics
Construction: Cellular Materials
Continuous Parallel Fiber Composites: Deformation and Strength
Continuous Parallel Fiber Composites: Fracture
Creep and Creep-fatigue of Metal-Matrix Composites
Creep of Concrete
Creep of Particle Reinforced Materials
Elastic Behavior of Cellular Solids
Elastic Behavior of Composites

Elasticity in Wood and Concrete: Hygromechanical Effects
Elastomers, Reinforcement of
Fatigue of Particle Reinforced Materials
Fiber Metal Laminates, Fatigue of
Fiber Strength
Laminates: Physical and Mechanical Behavior
Metal Matrix Composites with Roughly Equiaxed Reinforcements: Microstructure and Mechanical Behavior
Plastic Deformation of Cellular Materials
Polymer Composites, Friction and Wear of
Polymer Matrix Composites with Roughly Equiaxed Reinforcements: Microstructure and Mechanical Behavior
Reinforced Concrete

c) Characterization

Composites, Microstructure of: Quantitative Description
Composites: Interfaces
Mechanical Testing Methods of Fibers and Composites

VI – APPLICATIONS

Aircraft Materials
Automotive Body Materials
Automotive Chassis/Suspension Materials
Buildings: Plastics and Composites
Ceramic Matrix Composites: Applications
Cermets and Hardmetals
Composites for Biomedical Applications
Designing with Composites
Electronic Packaging: Conductive Adhesives
Electronic Packaging: Elastomer Conductive Polymers
Electronic Packaging: Heat Sink Materials
Liquid Crystalline Polymers, Dispersed
Materials Selection and Mechanical Design
Metal Matrix Composites: Applications
Optical Properties of Functional Hybrid Organic-Inorganic Nanocomposites
Polymer Matrix Composites: Applications
Polymer-based Nanocomposites by Sol-Gel Routes, Applications of
Polymer-Ceramic Nanocomposites: Catalysts
Radiation Effects in Carbon-Carbon Composites
Sporting Materials: Ski Equipment

This Page Intentionally Left Blank

COMPOSITE MATERIALS: OVERVIEW

The term "composite" originally arose in engineering when two or more materials were combined in order to rectify some shortcoming of a particularly useful component. For example, cannons that had barrels made of wood were bound with brass because a hollow cylinder of wood bursts easily under internal pressure. The early clipper sailing ships, which were said to be of composite construction, consisted of wood planking on iron frames, with the wood covered by copper plates to counter the attack of marine organisms on the wood.

For the purposes of this Encyclopedia, a composite material can be defined as a heterogeneous mixture of two or more homogeneous phases that have been bonded together. The phases can be of essentially any material class, including metal, polymer, ceramic, carbon, or even a fluid or void space: indeed, most highly porous materials are also composites. Provided the existence of the two phases is not easily distinguished with the naked eye, or provided the structural scale of the distribution of the two phases is far smaller than that of the structure that uses them, the resulting composite can itself be regarded as a homogeneous material with physical and mechanical properties of its own. Such materials are familiar. Many natural materials are composites, such as wood. Other examples are automobile tires, glass fiber-reinforced plastics (GRPs), the cemented carbides used as cutting tools, and paper—a composite consisting of cellulose fibers (sometimes with a filler, often clay). Paper is essentially a mat of cellulose fibers, combined with open pore space, with interfiber bonding being provided by hydrogen bonds where the fibers touch one another.

It is sometimes difficult to draw a distinction between a composite material and an engineered structure that contains more than one material and is designed to perform a particular function. The combination is usually referred to as a composite provided it has its own distinctive properties, such as being much tougher than any of its constituent elements alone, having a negative thermal expansion coefficient, or having some other property not clearly shown by any of the component materials.

The breaking strength of fibers of glass, of graphite, and of many whisker crystals, is much greater than that of bulk pieces of the same materials. In fact, it may be that all materials are at their strongest when in fiber form. In order to use the strength of such strong fibers, they must be stuck together in some way: for example, a rope is made up of fibers of hemp or flax, and indeed carbon or glass fibers are often twisted into tows. However, for maximum use, a matrix in which to embed such strong fibers is required, in order to provide a strong and stiff solid for engineering purposes. The properties of the matrix are usually chosen to be complementary to those of the fibers: for example, great toughness in a matrix complements the tensile strength of the fibers. The resulting combination may then achieve high strength and stiffness (owing to the fibers), and resistance to crack propagation (owing to the interaction between fibers and matrix).

The term "advanced composite" means specifically this combination of very strong and stiff fibers within a matrix designed to hold the fibers together. This type of composite combines the extreme strength and stiffness of the fibers and, owing to the presence of the matrix, shows much greater toughness than would otherwise be obtainable.

In many cases, the dimensions of one of the phases of a composite material are small, say between 10 nm and a few micrometers, and under these conditions that particular phase has properties rather different from those of the same material in the bulk form. When the microstructural scale falls in the range of a few nanometers, such a material is referred to as a "nanocomposite."

1. Making Composites

It is the act of artificially bonding distinct phases that makes a material a composite; this also creates a distinction between composite processing and bulk material processing. Many multiphased materials, such as high-carbon steel or porcelain, can indeed be produced by conventional bulk processing; however, in these the various phases are produced by internal phase transformation.

Most composite fabrication processes comprise two main steps: (i) separate production of the two phases to be combined, and (ii) combination of these two phases. The second step can be viewed as composite production in the stricter sense. Often, among the phases making a composite, one is distinguished as being the “reinforcement”: this is generally the strongest phase of the material. In most cases, the reinforcement is also geometrically distributed as discrete elements (fibers, particles) within the composite. The other phase, which bonds the reinforcement lements together and percolates throughout the composite, is then called the matrix. Although generally the weaker phase, the matrix influences markedly the properties of the composite. Composites are therefore classified according to the nature of their matrix: polymer, metal, and ceramic matrix composites, often designated PMCs, MMCs, and CMCs, respectively. In some cases, both phases of the composite percolate in all three dimensions throughout the material; the distinction between matrix and reinforcement then disappears to some extent. These composites are termed interpenetrating phase composites (IPCs).

Reinforcements are distinguished by their shape. They can be (i) long, essentially continuous fibers, (ii) short fibers, (iii) whiskers, which are small elongated single crystals of high perfection, or (iv) equiaxed particles of a wide variety of materials and forms. Processes by which the reinforcements are produced vary widely, and depend strongly on the reinforcement shape.

Fibers generally start as a viscous mass that is spun. This can be the material itself in molten form: glass is spun directly into fibers, as are metallic “melt-spun” ribbons. The fiber can also be spun in precursor form. This can be a slurry that is later dried and sintered, a polymer that is later pyrolyzed into a ceramic or into carbon, or pitch that can also be pyrolyzed into high-modulus carbon. Sometimes, fibers are not spun but gradually built by deposition onto a finer continuous fiber core. Thicker fibers result, which are generally called monofilaments. Thin metallic wires or filaments can also be produced in the solid state by drawing through a series of dies.

Short fibers can be spun directly if the process is sufficiently turbulent, or can start as continuous fiber that is chopped. Whiskers are grown using relatively complex processes that most often involve material transport from a vapor.

Equiaxed reinforcements are produced by many different processes: most ceramic particles are made by crushing bulk material, a process called comminution. Alternatively, equiaxed particles can be precipitated from a solution using sol–gel processing. Metallic particles can be solidified directly from the melt by atomization, can be produced from a precursor by precipitation or direct reduction, or can be cut from the bulk by machining.

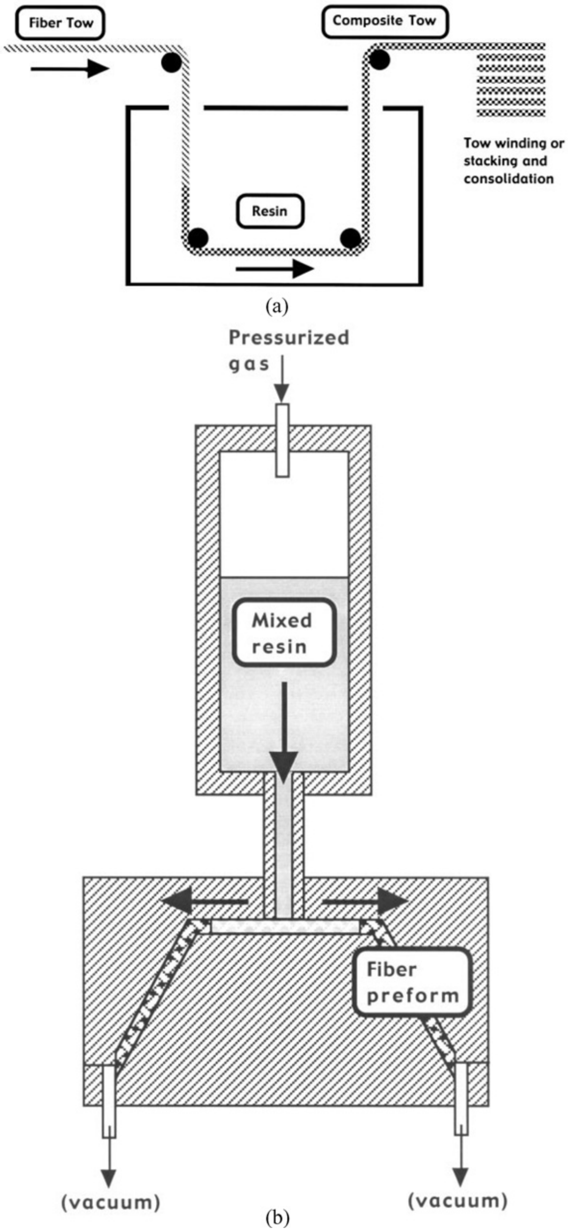
With such a wide spectrum of production routes, composite reinforcements include the cheapest materials available (sand qualifies as an equiaxed reinforcement, as do waste products such as flyash—in these cases, the reinforcement is often called a “filler”), as well as some of the most expensive structural materials on the market (e.g., boron monofilaments).

The process by which matrix and reinforcement are combined is dictated by the shape of the reinforcement and by the nature of the matrix. With polymer matrices, the composite can be produced at, or near, room temperature. This eases processing considerably: both directly because tooling is far simpler, and also indirectly because at high temperature one often encounters problems of chemical reactivity between reinforcement and matrix.

Polymer matrix composites with equiaxed reinforcements can be produced by simply stirring the particles into the matrix before curing the latter (with thermoset matrices) or at slightly above room temperature (with thermoplastic matrices). Short fibers can similarly be mixed into fluid polymers. Resulting viscous slurries can then be shaped in a closed mold or alternatively be coated onto a form; tooling required varies significantly with production volume and can be quite simple for small runs.

With continuous fiber reinforcements, processing becomes more complicated because of the need to control precisely how the fibers are handled and laid up. One method is to introduce the matrix into the tow, and then wind or stack the matrix-wetted tows to build up the composite. This route comprises filament winding and pultrusion processes used to produce composite cylinders or profiles (Fig. 1(a)). Alternatively, one can first lay up the dry fibers into a preform, and then infiltrate the fiber preform held in a mold with the matrix. These infiltration processes carry various names, one of the more usual being resin transfer molding (RTM) (Fig. 1(b)). For this latter class of processes, one can prepare the preforms by winding the fibers, or by weaving them into two-dimensional cloths or three-dimensional profiles. Woven fiber cloths can be stacked by hand, again allowing for relatively lightly tooled hand processing of composites.

With metal or ceramic matrices, processing becomes more complicated. This is the main reason why these two classes of composite materials lag far behind polymer matrix composites in engineering usage. One difficulty is that far higher temperatures must be employed, making handling and tooling more complex. Another difficulty stems from the fact that liquid metals and ceramics have far higher surface tensions than polymers, a result of the absence, in these matrices, of the weaker strictly physical bonds present in polymers. Although there are exceptions, molten inorganic matrices therefore tend not to wet spontaneously usual reinforcements. Here again, processes depend largely on the reinforcement geometry. Equiaxed reinforcement particles or short fibers can be stirred into the molten matrix. Alternatively, matrix and reinforcement powders can be blended and consolidated by conventional powder densification processes; this is, for example, how tungsten carbide tools are produced. With fibrous reinforcements, processes often start with the



(a)

(b)

Figure 1
Two processes used in the fabrication of fiber-reinforced polymer composites: (a) fiber tow impregnation and consolidation, often termed filament winding or pultrusion (depending on how impregnated tows are consolidated); (b) infiltration, often termed resin transfer molding (RTM).

production of a fiber preform that is subsequently infiltrated. Fiber-reinforced aluminum diesel engine pistons and SiC/SiC fiber composites are thus produced by liquid and vapor infiltration, respectively.

Since composites are generally produced by combination of two pre-existing phases, the interface between matrix and reinforcement is accessible during processing. Consequently, its nature and properties can be engineered to improve the composite or make it easier to produce. Depending on the desired composite properties, the interface can, for example, be made weak or strong, electrically insulating, or optically reflective. This is particularly important in brittle matrix fiber composites that are toughened by the fiber pull-out mechanism, in which the interface must be sufficiently weak to cause fiber/matrix debonding once the interface is reached by a crack. Continuous fibers destined to be combined with a polymer matrix are thus generally coated with a thin layer of "sizing," which protects the dry fibers against mutual rubbing and produces an optimal bond between matrix and reinforcement. In metal and ceramic matrix composites, interface engineering is often conducted with the aim of protecting the reinforcement from chemical attack by the matrix, or to promote wetting of the reinforcement by the matrix. To this end, coatings of many different types have been employed. Perhaps the most sophisticated example is that of layered carbon/SiC coatings applied to SiC monofilaments for combination with titanium-based matrices: the carbon layer makes the interface weak, while the SiC protects the carbon layer from the titanium.

Sometimes, the two main steps of composite processing, namely (i) reinforcement and/or matrix production and (ii) combination of the two phases, occur simultaneously or in inverse order. Such processes are generally called "*in situ*" processes, because one of the phases, generally the reinforcement, is created by chemical reaction within a bulk material that is a precursor to the composite. Examples include the directed oxidation of aluminum to produce its oxide toughened by channels of remanent metal, or the consolidation of powder blends of different phases that react with one another to form, for example, an intermetallic and a ceramic. Many highly porous materials are also produced using in-situ processes; an example is with polyurethane foam, which is produced by reactive gas evolution within a viscous liquid precursor of the material containing surface-active additions that stabilize thin cell walls against rupture. Although often essentially "bulk" in nature, *in situ* composite processes can generally be distinguished from alloying or other conventional bulk material processes by the fact that they involve phase formation by chemical reaction instead of phase transformation. Still, the term "*in situ* composite" has often been used by analogy to designate some conventionally processed bulk materials, either because of the fibrous microstructures produced (*in situ* eutectic composites fall in this class), or because the two phases are widely different, such that the material behavior is similar to that of certain composites with equiaxed reinforcements (metal alloys containing brittle nonmetallic second phases have thus at times been termed composites).

Cost is of course an important consideration in composite processing, as it ultimately determines whether composites will be used in practice. Many composites can be molded and shaped into highly complex shapes: this can make these materials cost-competitive despite higher material prices if one composite part replaces a structure that would otherwise be assembled by welding or riveting several metallic components. Tooling is a particularly important issue in composite mass production. With composites, tooling can suffer because of wear (e.g., ceramic particle-reinforced aluminum or PEEK/carbon composites tend to erode dies rapidly), or can be critical because, unlike a metal that can be bent back to shape, polymer composite components are systematically rejected if tolerances are not respected. With composites, recycling is also a complex issue. The material can sometimes be reused; however, other methods of recycling may be opportune, such as separating matrix and reinforcement before recycling each separately. The competitiveness of composites thus hinges on many issues other than the intrinsic cost or performance of the material itself.

2. Predicting Properties of Composites

An important question, both for engineering design of a composite material and for scientific understanding of its properties, is that of how the overall properties of the composite depend on those of the individual constituents.

Properties of a composite material can be considered under two headings. (a) those that depend solely upon the geometrical arrangement of the phases and their respective volume fractions, and not at all upon the dimensions of the components, and (b) those that depend on structural factors such as periodicity of arrangement or the sizes of the pieces of the two or more component phases. Indeed, in many instances the properties of phases making the composite start changing when the smallest dimension of the phase (the diameter of a fiber, the thickness of the cell wall in a foam) falls below a certain dimension. The plastic deformation of metals is for example strongly altered when metal elements are less than 1 μm thin, and the band gap of semiconductors starts changing radically when their size falls below roughly 10 nm.

2.1 Properties Determined by Geometry—Additive Properties

The geometrical arrangement of the phases in a composite material can be often described in simple terms (Fig. 2); these can also be quantified using principles of image analysis. Modern computed tomography gives access to a full three-dimensional description of the material microstructure if there is sufficient X-ray contrast between the phases. Such is obviously the case with highly porous materials such as foams and microcellular materials, for which the technique is proving very powerful indeed (Baruchel *et al.* 2000, Maire *et al.* 2003).

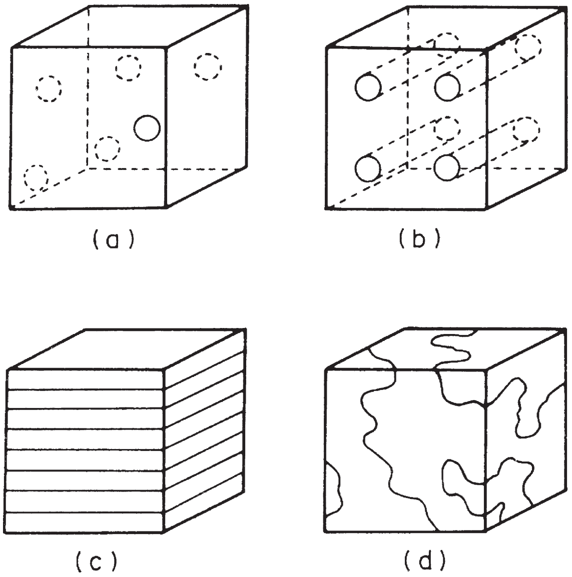


Figure 2
Composite geometries: (a) random dispersion of spheres in a continuous matrix; (b) regular array of aligned filaments; (c) continuous laminas; (d) irregular geometry (after Hale 1976).

The simplest physical property of a composite, namely its density, ρ_c, is given by the volume-weighted average of the densities of the components:

$$\rho_c = \rho_1 V_1 + \rho_2 V_2 \tag{1}$$

where V is the volume fraction and subscripts 1 and 2 refer to the components. If there are no other phases (including voids), $V_1 + V_2 = 1$. The volume fraction of each phase in the composite can thus be computed by measuring its density knowing that of its components; since the precision of this measurement increases with increasing density contrast between the phases, it is particularly simple and trustworthy for highly porous materials such as a foam.

More complicated physical properties, e.g., those described by a second-rank tensor, relate two vectors: a solenoidal vector and an irrotational vector (as with magnetic or electric susceptibility), or else a flux vector and the gradient of a scalar function (as with diffusivity and electric and thermal conductivity). The relationships derived for these properties for a composite material in terms of the same property in each of the submaterials or components are all formally identical. They have usually been discussed in terms of the dielectric constant, ε, and, because of this, the general attempt to predict the overall physical properties of a multiphase composite from the same properties of the individual phases is sometimes called the dielectric problem. Attempts to devise exact, generally applicable theoretical expressions for the dielectric constants of mixtures of unknown phase geometry are futile. However, the dielectric constant of a mixture of two phases, ε_c, must lie between certain limits, whatever the geometry. If some knowledge of the phase geometry is available, still closer bounds can be set on the value of ε_c; the same is true of other properties. There is, therefore, a good deal of utility in the search for bounds on the properties. For a few special properties, exact solutions can be obtained. Approximate solutions are also often obtained for two-phase composites when the volume fraction of one of the components is low. Such solutions are often extended to nondilute situations by use of a self-consistent scheme, in which each particle of the minor component is assumed to be surrounded not by the other component but by the average composite material. For example, as given by Hashin and Shtrikman (1962), the effective dielectric constant of an isotropic composite must lie in the range

$$\varepsilon_1 + \frac{V_2}{\dfrac{1}{\varepsilon_2 - \varepsilon_1} + \dfrac{V_1}{3\varepsilon_1}} \leqslant \varepsilon_c \leqslant \varepsilon_2 + \frac{V_1}{\dfrac{1}{\varepsilon_1 - \varepsilon_2} + \dfrac{V_2}{3\varepsilon_2}} \qquad (2)$$

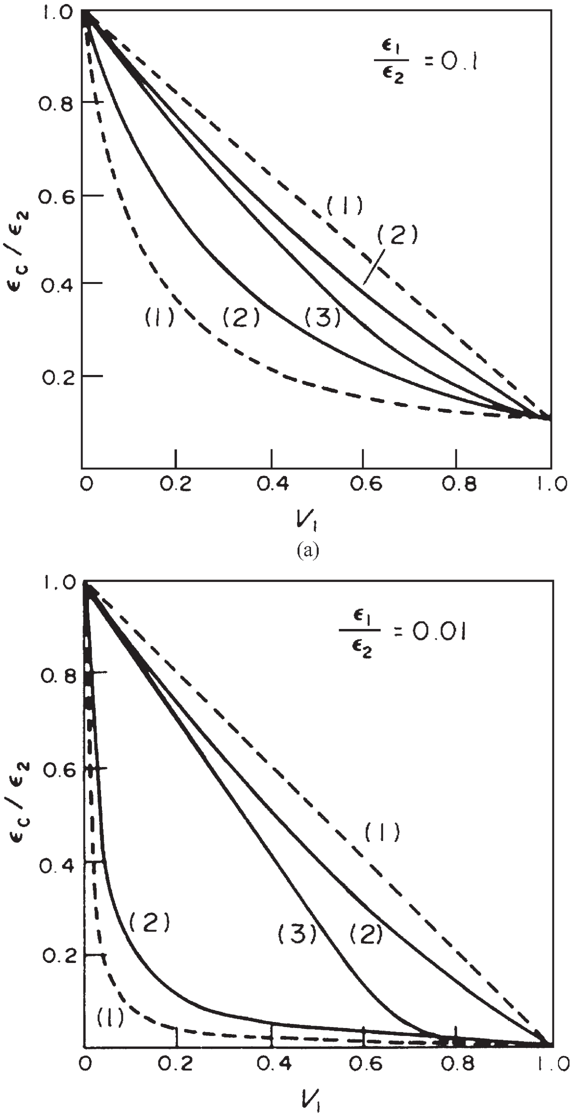
where $\varepsilon_2 > \varepsilon_1$. These are the best possible bounds obtainable for the dielectric constant of an isotropic two-phase material if no structural information, apart from the volume fractions, is available.

Figure 3 compares the predictions of those bounds for the cases $\varepsilon_1/\varepsilon_2 = 0.1$ and $\varepsilon_1/\varepsilon_2 = 0.01$ with the most elementary bounds that can be rigorously derived for an isotropic composite. These elementary bounds correspond to simple Voigt (polarization in parallel) and Reuss (polarization in series) estimates. The Voigt estimate is of the form of Eqn. (1), while the Reuss estimate gives the composite dielectric constant as

$$1/\varepsilon_c = V_1/\varepsilon_1 + V_2/\varepsilon_2 \qquad (3)$$

If ε_1 and ε_2 do not differ greatly, the bounds given by Eqn. (2) are very close together, but for $\varepsilon_1/\varepsilon_2 = 0.01$ they are far apart (such as in Fig. 3). Therefore, in many practical cases, better bounds are required.

In the above discussion, the properties of the individual components or submaterials within the composite are assumed to be identical with the same properties measured in a piece of the submaterial outside the composite. In addition, the composite is taken to be statistically homogeneous in the sense that, if small elements of the material were extracted, these would have the same physical properties as the whole sample. In addition, the implicit assumptions are made that there are no voids, that space-charge effects and polarization are absent (in the case of electrical conductivity), and that there is no discontinuity in temperature at the interface (in the case of heat flux).



(a)

(b)

Figure 3
Dielectric constant of composite materials: (a) $\varepsilon_1/\varepsilon_2 = 0.1$ and (b) $\varepsilon_1/\varepsilon_2 = 0.01$. (1) Bounds from Eqns (1) and (3); (2) Hashin and Shtrickman bounds for arbitrary geometry (Eqn. (2)); (3) self-consistent approximation for spheres in a continuous matrix (after Hale 1976).

Thermal expansion coefficients of composites involve both the elastic constants and the thermal expansion coefficients of the individual phases. An important function of reinforcing fillers and fibers in plastics is the reduction and control of thermal expansion. For example, with dental filling materials, a difference in thermal expansion between the filling material and the tooth substance can lead to a marginal gap, and hence composite filling materials are designed to have a thermal expansion coefficient very close to that of the tooth substance.

With some components, especially at low fiber volume fractions, the transverse thermal expansion coefficient of, for example, a glass-fiber epoxy-resin composite, can be greater than that of the matrix. This effect is particularly noticeable with fibers of high modulus and low expansion coefficient (e.g., boron or carbon) in a low-modulus matrix.

Fibers are often used in laminated arrangements because properties parallel to the fibers are very different from properties perpendicular to the fibers. The effective in-plane thermal expansion coefficients for angle-ply laminates (in which the fibers are arranged at plus and minus an angle θ to a particular direction) show that in such laminates a scissoring or lazy-tongs type of action can occur, and, with appropriate values of θ, can lead to a zero or even negative thermal expansion in one direction.

2.2 Special Property Combinations—Product Properties

Composite materials can, in principle, be thought of as materials that produce properties unobtainable in a single material. For example, by combining a piezoelectric material with a material showing magnetostriction, the composite should allow a magnetoelectric effect: i.e., an applied magnetic field would induce an electric dipole moment; or a material could be produced in which an applied magnetic field produced optical birefringence, by coupling a material that shows strain-induced birefringence with one showing magnetostriction.

This idea led van Suchtelen (1972) to classify such effects as product properties of composites and so now there are considered to be two different types of physical property of composite materials.

The first is the type discussed in Sect. 2.1. These are sum or additive properties where the composite property is related to that same property of each of the components and so depends on the geometry of arrangement of the two components. The geometry of arrangement includes of course the volume fraction. Examples are elastic stiffness, relating applied stress and measured strain, or electrical resistance, relating applied electric field and measured current density, or the simple example of mass density. The value of the physical property of the composite in general lies between those of the components.

There is a subclass of additive properties in which the value of the property of the composite can lie well outside the bounds set by the values of the property of the components. Examples are Poisson's ratio involving the ratio of two compliance coefficients under the action of a single applied stress, or acoustic wave velocity, which depends on the ratio of elastic modulus to density. Here, because the elastic modulus and the density can follow quite different variations with volume fraction, e.g., elastic modulus following Eqn. (3) and density following Eqn. (1), the acoustic wave velocity of the composite can lie well outside the value for either component. A further, but less obvious, example is the thermal expansion coefficient, which depends on the thermal expansion coefficients and the elastic constants of the two phases.

Product properties are those cited at the beginning of this section: the property of the composite depends specifically on the interaction between its components. Any physical property can be considered as the action of a physical quantity X resulting in physical quantity Y giving the X–Y effect. Product properties are those in which an X–Y effect in submaterial 1 produces a Y–Z effect in submaterial 2, producing in the composite an X–Z effect. A good example is that of a magnetoelectric effect in a composite material having one magnetostrictive and one piezoelectric phase. Application of a magnetic field produces a change in shape in the magnetostrictive phase that then stresses the piezoelectric phase and hence generates an electric field. In this case the coupling is mechanical but the coupling could also be electrical, optical, magnetic, thermal, or chemical (van Suchtelen 1972).

Viewed in this way, sum properties are those in which an X–Y effect in submaterial 1 and the same X–Y effect in submaterial 2 combine to give an X–Y effect in the composite.

Table 1 classifies some physical properties or phenomena according to the input–output parameters X and Y. A small selection of possible product properties is given in Table 2.

2.3 Properties Dependent on Phase Dimensions and Structural Periodicity

In Sects. 2.1 and 2.2 the physical properties of the components or submaterials are assumed to be unaltered by the incorporation of the component into the composite. This is usually not the case. The effects of phase changes or chemical reactions during fabrication are of importance in a wide range of composites. Excessive shrinkage of one component can result in high internal stresses, which may lead to premature failure or even preclude successful fabrication. Matrix shrinkage also has an important indirect effect on the mechanical behavior of fiber

Table 1

Matrix classification of some physical properties or phenomena in materials according to the type of input and output parameters (after van Suchtelen 1972)

	Input parameter (X)					
Output parameter (Y)	Mechanical (force/deformation) (1)	Magnetic (field/polarization) (2)	Electrical (field/polarization, current) (3)	Optical and particle radiation (light or particle flux) (4)	Thermal (temperature, temperature gradient, heat current) (5)	Chemical (chemical composition, chemical composition gradient) (6)
Mechanical (force/deformation) (1)	elasticity	magnetostriction magnetoviscosity (suspension)	electrostriction Kirkendall effect electroviscosity (suspension) indirect: thermal expansion		thermal expansion	osmotic pressure
Magnetic (field/polarization) (2)	piezomagnetism	magnetic susceptibility	superconductors galvanic deposition of ferromagnetic layer direct generation of magnetic field	photomagnetic effect	thermomagnetism ferromagnetic material at $T \simeq T_c$ (+magnetic field)[a]	dependence of T_c on ferromagnetic composition
Electrical (field/polarization, current) (3)	piezoelectricity piezoresistivity	magnetoresistance (+electric current)[a] Hall effect (+electric current)[a] ac resonance induction of voltage	dielectric constant, dielectric polarization Hall effect (+magnetic field)[a]	photoconductivity photoemission photoelectromagnetic effect (+magnetic field)[a] ionization	thermoelectricity ferroelectrics at $T \simeq T_c$ (+electric field)[a] temperature-dependent resistivity (+electric current)[a]	dependence of T_c on ferroelectric composition
Optical and particle radiation (light or particle flux) (4)	stress birefringence triboluminescence	Faraday effect magnetooptic Kerr effect deflection of charged particles	electroluminescence laser junctions refractive index Kerr effect absorption by galvanic deposits cold emission of electrons	refractive index fluorescence scintillation color-center activation	thermoluminescence	chemoluminescence
Thermal (temperature, temperature gradient, heat current) (5)	heat of transition of pressure-induced phase transition piezoresistivity and Joule heating	adiabatic demagnetization Nernst-Ettingshausen temperature gradient effect (+electric current)[a] magnetoresistance effect + Joule heating (+electric field)[a]	dissipation in resistance Peltier effect Nernst-Ettingshausen temperature gradient effect (+magnetic field)[a]	absorption	thermal conductivity	reaction heat
Chemical (chemical composition, chemical composition gradient) (6)	pressure-induced phase transition		electromigration galvanic deposition	light or particle stimulated reactions (photosensitive layers)	Soret effect (temperature gradient)[a] phase transition change of chemical equilibrium	

[a] indicates the parameter in parenthesis is essential as a second input.

Table 2
Product properties of composite materials (after van Suchtelen 1972)

X–Y–Z (Table 1)	Property of phase 1 (*X–Y*)	Property of phase 2 (*Y–Z*)	Produce property (*X–Z*)
1–2–3	Piezomagnetism	Magnetoresistance	Piezoresistance Phonon drag
1–2–4	Piezomagnetism	Faraday effect	Rotation of polarization by mechanical deformation
1–3–4	Piezoelectricity	Electroluminescence	Piezoluminescence
1–3–4	Piezoelectricity	Kerr effect	Rotation of polarization by mechanical deformation
2–1–3	Magnetostriction	Piezoelectricity	Magnetoelectric effect
2–1–3	Magnetostriction	Piezoresistance	Magnetoresistance Spin-wave interaction
2–5–3	Nernst–Ettingshausen effect	Seebeck effect	Quasi-Hall effect
2–1–4	Magnetostriction	Stress-induced birefringence	Magnetically induced birefringence
3–1–2	Electrostriction	Piezomagnetism	Electromagnetic effect
3–1–3	Electrostriction	Piezoresistivity	Coupling between resistivity and electric field (negative differential resistance, quasi-Gunn effect)
3–4–3	Electroluminescence	Photoconductivity	
3–1–4	Electrostriction	Stress-induced birefringence	Electrically induced birefringence light modulation
4–2–1	Photomagnetic effect	Magnetostriction	Photostriction
4–3–1	Photoconductivity	Electrostriction	Photostriction
4–3–4	Photoconductivity	Electroluminescence	Wavelength changer (e.g., infrared-visible)
4–4–3	Scintillation	Photoconductivity	Radiation-induced conductivity (detectors)
4–4–4	Scintillation, fluorescence	Fluorescence	Radiation detectors, two-stage fluorescence

composites, since the resulting internal stresses can determine the frictional forces at the fiber/matrix interface.

The dimensions and periodicity of a composite structure also have an important effect on the properties when they become comparable with, for example, the wavelength of incident radiation, the size of a magnetic domain, or the thickness of the space-charge layer at an interface. The dimensions are also particularly important when the properties of a material depend on the presence of defects of a particular size (e.g., cracks). The strength of a brittle solid containing a crack depends on the square root of the length of the crack. Since very small particles cannot contain long cracks, and because the surface region of most solids behaves differently from the interior (a small fiber or sphere contains proportionately more surface material than a large piece), the breaking strength depends on the dimensions of the piece tested. This is particularly important when considering mechanical strength and toughness of composites.

Many of the optical effects that depend on structural dimensions also depend in a very complex way on other factors, and optical effects produced by dispersion of a second phase within a material are often used to determine the distribution of that phase (e.g., determination of the molecular weights of polymers or of the size of crystal nuclei in glasses), rather than regarding the material as a composite with special optical properties. However, a composite material containing aligned elongated particles of an optically isotropic material in an optically isotropic matrix may exhibit double refraction as a straightforward consequence of the relationship between dielectric constant and refractive index.

In a ferromagnetic material at temperatures below the Curie point, the electronic magnetic moments are aligned within small contiguous regions—the magnetic domains—within each of which the local magnetization is saturated. Magnets of very high coercivity can be obtained by using very fine particles: a particle with a diameter of between 10^{-1} μm and 10^{-3} μm would be magnetized almost to saturation since the formation of a flux-closure configuration would be energetically unfavorable. Magnetization reversal cannot take place in a sufficiently small single-domain particle by boundary displacement; it must occur by a single jump against the magnetocrystalline anisotropy and the shape factor. Fine-particle permanent magnets of the rare-earth alloys such as $\mathrm{SmCo_5}$, enveloped in an inert matrix such as tin, have been developed that have superior properties.

3. Advanced Fiber Composites

The best-known and striking example of a modern composite material is, as mentioned above, the fiber composite, designed for high strength, high stiffness, and low weight.

The stiffest materials are ceramics composed of elements from the first rows of the periodic table (B, C, SiC, Al_2O_3, etc.). This is because these elements contain the highest possible packing of strong covalent bonds directed in three dimensions. Such materials also each have a high melting point, a low coefficient of thermal expansion, and a low density. All of these are very desirable engineering properties.

In all structural applications, the weight of the structure necessary to bear a given load and to do so with a minimum elastic deflection is important, and so strong fibers are compared on the basis of specific modulus and specific strength, i.e., modulus or strength divided by density, respectively. When this is done the stiffest materials on a weight-for-weight basis are the plastics (PET, PBT, Kevlar) and the glasses. The metals make a very poor showing on either specific stiffness or specific strength. However, fibers cannot be used directly and must be bound together with a matrix.

A set of aligned fibers in a pliant matrix is only useful as an engineering material under direct tensile forces parallel to the fibers. The salient properties are represented by those of a pocket handkerchief, which is stiff in the direction of the warp and weft (parallel to the fibers) but shears easily parallel to these. Hence the properties of fiber composites are very directional, being strong and stiff parallel to the fibers, but rather weak in shear parallel to the fibers (because this property depends on the shear properties of the matrix), and often very weak indeed in tension perpendicular to the fibers. In order to overcome this, the fibers can be packed along several directions; patterns exist that permit this in three dimensions, some of which actually produce fully isotropic structures (Parkhouse and Kelly 1998). In practice, as advanced fiber composites are often used to produce relatively thin shell structures (such as the hull of a sailboat, the body of an automobile or the wing of an airplane), the fibers are arranged in planar sheets called laminae, each containing parallel fibers, and these are stuck together so as to provide a more isotropic material with a high volume loading of fibers.

Alternatively, the fibers may be randomly arranged in a plane or in three dimensions; such arrangements limit the obtainable fiber packing density (Parkhouse and Kelly, 1995). Fibers are also often woven into mats before incorporation into the composite because this aids the handling of fiber arrays.

When carbon- or other fiber-reinforced plastics are made with woven fabric rather than nonwoven material, distortion of the load-carrying fibers parallel to the applied stress reduces the tensile strength and reduces stiffness and toughness. When, however, woven fabric is oriented at 45° to the load direction these properties compare favorably with those for nonwoven $\pm 45°$ material. Indeed, the residual strengths after impact can be greater.

The anisotropy of the properties of advanced fiber composites represents a unique feature in engineering design using advanced composites. The anisotropy can be varied; in fact it is possible now to place the fibers precisely where they are needed in a structure to bear the loads. It is also possible, for instance, to vary the amount and stacking of the fibers in various directions along the length of a beam so as to vary the stiffness and torsional rigidity of the material.

It would be impossible to make use of the inherent advantages of anisotropic materials such as carbon fiber reinforced plastics (CFRPs) without the invention of computer-aided methods. The reiterations and the number of variables involved require computer programs both to relate the properties of the individual laminae to the properties of the fibers and to take into account the interactions in term of bending–twisting coupling between the various laminae. Initially, of course, these complications were viewed as a great disadvantage of the material, since the engineer was not happy with anisotropy. Now the problem of how to use it and make a virtue, not so much out of necessity, but of a supposed vice, has been solved. The Grumman X29 aircraft makes use of bending–twisting coupling so that, as the loads on an aircraft wing increase, the structure can deform but remain tuned to the aerodynamic requirements. The material possesses almost a form of "gearing", so that it changes its shape automatically without the need for sensors and levers.

The very strong and stiff fibers comprise asbestos (a naturally occurring fiber), boron, and carbon of varying degrees of graphitization, made in principle from a number of starting materials but in practice mainly two: PAN (polyacrylonitrile) and pitch. Inorganic glasses based on silica yield many fibers. Those made in the form of continuous filaments are most important for composite materials. Many linear polymers, such as polyparabenzamide or simple polyethylene, can be processed to form very stiff and strong fibers. Ceramic fibers, of Al_2O_3, mullite, and silicon carbide and nitride, are important for high-temperature use, either as a reinforcement for metals or ceramics, or in filtration and other applications because of their chemical inertness. Additionally, fibers can be sensors. For example, if made of shape memory alloys (SMAs) the composite material can then react to its environment, by sensing or by deforming under the influence of variations in its mechanical or thermal environment. Such composites have been termed "smart" or "intelligent".

Strong fibers, whether they are stiff or not, also have the great advantage of restraining cracking in what are called brittle matrices. It is this use of fibers

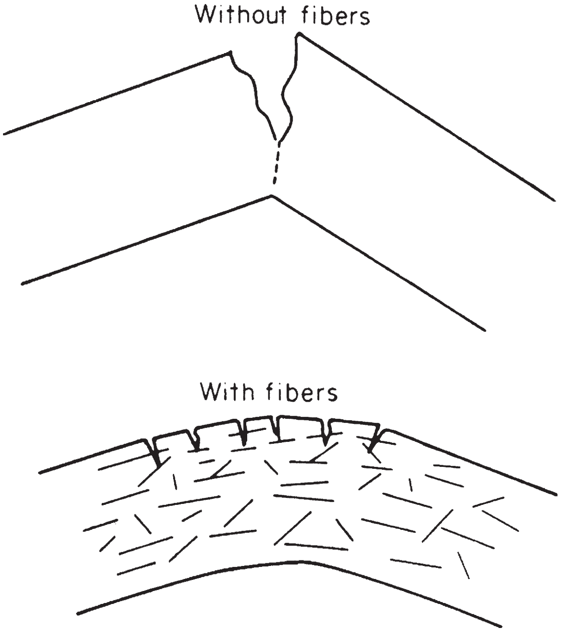


Figure 4
Demonstration of how cracks are prevented from running in a brittle material because of fibers in their path.

that is often referred to when showing how the ancients employed composite materials. Quotations from the Bible are made, such as that from Exodus, Chapter 5, verses 6 *et seq.*, which refer to the difficulty of making bricks without straw. Fibers restrain cracking because they bridge the cracks. The principle is very simple. It has been used in asbestos-reinforced cement and the principle of reinforced concrete is not too far from the same idea. The principle is illustrated in Fig. 4. A crack passing through a brittle material may not enter the fibers but leaves a crack straddled by fibers so that a material that would normally have broken with a single crack now has to be cracked in a large number of places before the fibers themselves fail or pull out. This is used in fiber-reinforced cement, which is based on the same principles as reinforced concrete, but since the fibers are very much finer than reinforcing bars and are often not visible to the naked eye, a homogeneous material is effectively produced that can show quasi-plastic deformation.

4. *Composites with Short Reinforcements*

Many composite materials are reinforced not with continuous fibers but with more or less equiaxed reinforcements, such as particles or short fibers. In these, the transfer of applied stress from matrix to reinforcement is less effective than it is with a continuous fiber reinforcement; the properties of such composites are, therefore, dominated to a far higher degree by those of the matrix. Still, equiaxed reinforcements can significantly improve the properties of the matrix, and generally do so at far lower cost than with fibers. The reinforcement cost can even be so low that it is mainly added to the matrix because it is cheap: mineral particulate fillers are at times incorporated in polymers or low-cost construction materials chiefly for this reason. Generally, however, one adds the reinforcement in order to improve matrix properties.

Friction and wear resistance can, for instance, be much improved with equiaxed reinforcements: tribological behavior being largely determined by superficial material only, it matters less if the reinforcement is fibrous or not. Particles of hard ceramic added to polymers or metals thus produce very good improvements in wear and abrasion resistance, and do so about as well as fibers would. An example is provided by SiC particle-reinforced aluminum. This composite is used to produce high-performance brake disks, which are strong, light, excellent heat conductors, and sufficiently refractory. Cemented carbide cutting tool materials provide another example of a composite with an equiaxed reinforcement that is largely designed to resist abrasion. Composites containing lubricating reinforcements (generally graphite) can also be produced to reduce dry friction.

Although less spectacularly than with a fiber, the toughness of a brittle matrix can also be improved with the addition of whiskers, short fibers, or equiaxed particles. This can be accomplished by several mechanisms. The reinforcement can be a far tougher phase than the matrix, which (provided the crack does not meander its way between particles) prevents crack propagation by creating bridges that pull together matrix crack faces. Elastomeric additions to polymers and metal islands or channels within a ceramic matrix play this role. Another effective toughening mechanism achieved with equiaxed reinforcements is obtained in monolithic ceramics with the addition of zirconia particles that present a tetragonal to monoclinic phase transformation. More simply, equiaxed reinforcements can also toughen a brittle matrix by creating a site for localized deflection of cracks along their path.

With equiaxed reinforcements, the composite behavior is more like that of conventional bulk materials. The science of multiphase metal alloys or ceramics can thus be used to a large extent towards understanding the behavior of this class of composites. However, composites are often combinations of highly dissimilar materials, such as a metal and a ceramic, and this exacerbates the complexity of their interaction. Another complication in analysis is, as discussed above, that the properties of each phase may be different in the composite than in the same phase when it is in bulk form. Its microstructure may differ (e.g., if the reinforcement nucleates matrix crystallization), as may its strength, either because the defect population differs or because the agents of deformation, such as dislocations, move and interact differently in the narrow space between two matrix/reinforcement interfaces. These effects increase in importance as the scale of the microstructure of the composite is refined.

Sometimes it is their greater similarity and compatibility with isotropic bulk material that makes composites reinforced with short or equiaxed reinforcements attractive. If isotropic properties are desired, for reasons of design or for ease of processing, equiaxed reinforcements are thus useful. Also, particulate composites may be preferred over fiber composites because they are far easier to join, particularly with bulk material components. For example, particle-reinforced metals can be welded, which enables their economical use in bicycle frames or automotive driveshafts.

5. *Porous and Cellular Materials*

A special class of composite materials are those in which a material is combined with empty pore space. Indeed, although porosity is a defect in many materials, causing premature failure in incompletely sintered ceramics or poorly fed metal castings for example, sometimes it is desirable.

There are various reasons why one might want to introduce pores in a material. One is to create a light and soft material: polyurethane foam is a commonplace example that is used in applications requiring a low modulus or high compressibility: furniture, sports shoes are examples. Another is thermal insulation. Indeed, small empty pores are about as good an insulator as can be found; hence materials designed to resist heat conduction are generally very porous: styrofoam and glasswool are examples. Filters, sieves, catalysts, fuel cell electrodes, sound absorbers are all applications that call for highly porous materials. Also, many of Nature's materials are microcellular: wood and bone for example; artificial tissue and bone replacement materials are hence also often porous.

Porous materials are often composites in the sense that pores are deliberately introduced into the base material; also, these are engineered in that the pore size and shape are purposefully controlled. Many methods exist to produce highly porous materials. One is foaming: if a reactive gas former is mixed into a viscous material, or if a material is highly supersaturated with gas, bubbles can be made to nucleate and grow. If these are sufficiently stable, the bubbles can grow to many times the volume of the material in which they have nucleated. Bubble stability is promoted by a "disjoining" pressure that pushes apart two liquid/gas interfaces as they approach one another; reducing the surface tension of the liquid

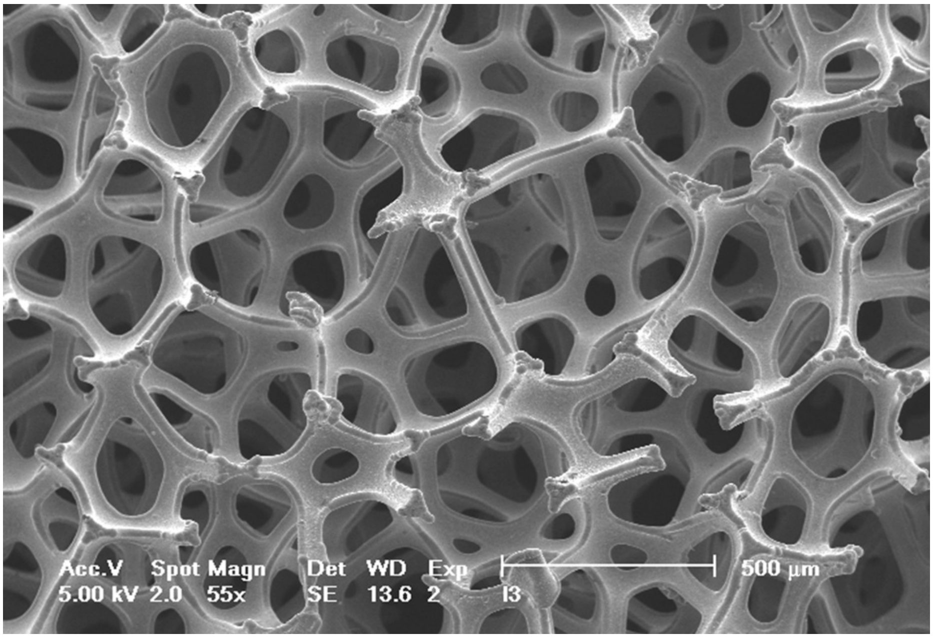


Figure 5
Highly porous microcellular nickel material produced by electrodeposition onto a metallized open-pore polyurethane foam. (Sample produced by Nitech S.A., France; sample kindly provided by Dr. Croset).

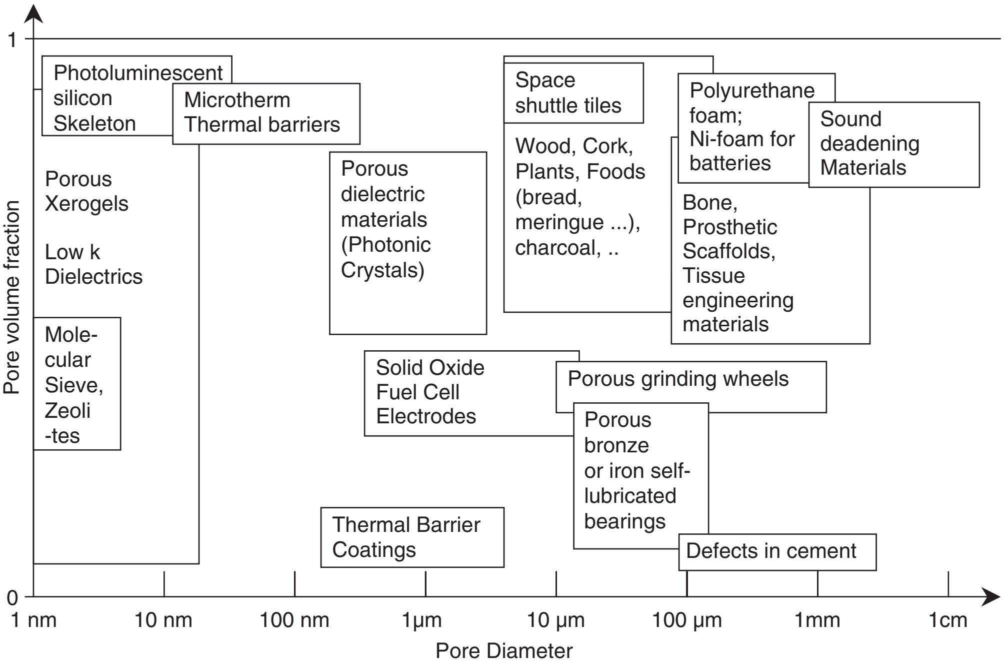


Figure 6
A map giving roughly the scale and level of porosity of various porous materials. Illustrations for several of these can be found in (Kelly, Clyne, Bonfield, and Willoughby, 2006).

also helps: these are engineered to prevent bubble walls from becoming unstable and bursting. After solidification of the material if it is liquid, the result is a highly porous solid foam, in which pores are individually isolated, or ‘‘closed’’.

In another class of highly porous materials, the pores are ‘‘open’’ or interconnected, and hence accessible to an external fluid: filters or paper are examples. Open-pore materials can be made by bursting the walls between bubbles in a foam, by partially bonding or sintering a packed bed of particles or short fibers, by infiltrating a fine-scale place holder material that is later removed, or by chemical attack of alloyed materials. Figure 5 gives an example of an open-pore microcellular material: this is a nickel battery electrode material that was created by electrodepositing the metal onto an open-pore polyurethane foam.

Polymers, metals, ceramics, carbon and biomaterials have all been used to produce highly porous or microcellular materials, featuring both open and closed pores (Gibson and Ashby, 1997; Schüth *et al.* 2002; Kelly *et al.* 2006). In some cases, the porous material is itself a composite: solid oxide fuel cell electrodes, generally a highly porous nickel/ceramic composite, are one example.

An important parameter in most porous materials is the pore size; essentially, one can distinguish between microporous and nanoporous materials. In polyurethane foams, highly porous metals, or many natural microcellular materials such as wood and bone, the pores are many micrometers in diameter. In other porous materials, individual pores are only a few nanometers in diameter: natural zeolites are one example; nanoporous silicon is another. Such small pores can have interesting properties. For example, when solid struts in nanoporous silicon are only a few nanometers across they become photoluminescent. Thermal insulators also become particularly efficient if the pores are less than 100 nm in diameter. Figure 6 illustrates the wide range of pore scale and ‘‘relative density’’ (i.e., the solid phase volume fraction in the porous material, equal to one minus the pore volume fraction) that are found in today’s highly porous materials.

6. *Conclusion*

The term ‘‘composite material’’ collects under one heading a very large and now ubiquitous class of materials. Their common and defining trait is that composites are combinations of phases that are bonded together instead of being created by phase transformation within bulk materials. Composites can therefore contain widely different and interacting materials, combined at a scale and with a microstructural architecture that are controlled with considerable latitude. This enables the generation of materials with special and controlled properties, making composites unique among materials. Their main limitations are generally their cost, as well as the challenges they pose in design, recycling, shaping, and assembly. These limitations are offset by significant and often spectacular gains in performance, such that the use of composite materials, already very high, is still increasing rapidly.

Bibliography

Bader M G, Kedward K K, Sawada Y (eds.) 2000 Design and applications. *Comprehensive Composite Materials*, Vol. 6. Pergamon, Oxford, UK

Baruchel, J, Buffière JY, Maire E, Merle P and Peix G (eds.) 2000, *X-ray Tomography in Materials Science*. Hermès Science Publications: Paris, France. 204 pages.

Bunsell A R (ed.) 1988 *Fibre Reinforcements for Composite Materials*. Elsevier, Amsterdam

Bunsell A R, Kelly A, Massiah A (eds.) 1993 Developments in the science and technology of composite materials. *Proc. 6th European Conf. on Composite Materials*. Woodhead, Cambridge, UK

Carlsson L (ed.) 2000 Test methods, nondestructive evaluation, and smart composites. *Comprehensive Composite Materials,* Vol. 5. Pergamon, Oxford, UK

Chawla K K 1998 *Composite Materials, Science and Engineering*, 2nd edn. Springer, New York

Christensen R M 1979 *Mechanics of Composite Materials*. Wiley, New York

Chou T W 1992 *Microstructural Design of Fiber Composites*. Cambridge University Press, Cambridge, UK

Chou T W (ed.) 2000 Fiber reinforcements and general theory of composites. *Comprehensive Composite Materials,* Vol. 1. Pergamon, Oxford, UK

Clyne T W (ed.) 2000 Metal matrix composites. *Comprehensive Composite Materials,* Vol. 3. Pergamon, Oxford, UK

Clyne T W, Withers P J 1993 *Introduction to Metal Matrix Composites*. Cambridge University Press, Cambridge, UK

Cogswell F N 1992 *Thermoplastic Aromatic Polymer Composites*. Butterworth Heinemann, Oxford, UK

Gerstle Jr. F P 1986 Composites. In: Marle H F , Bikales N M, Overberger C G, Menges G (eds.) *Encyclopedia of Polymer Science and Engineering,* 2nd edn., Vol. 3. Wiley, New York, pp. 776–820

Gibson L J, Ashby M F 1997 *Cellular Solids - Structure and Properties*, Second Ed., Cambridge Solid State Science Series, Cambridge.

Gutowski T G (ed.) 1997 *Advanced Composite Manufacturing*. Wiley, New York

Hale D K 1976 The physical properties of composite materials. *J. Mater. Sci.* **11**, 2105–41

Hannant D J 1978 *Fiber Cements and Concretes*. Wiley, New York

Hashin Z, Shtrikman S 1962 A variational approach to the theory of the effective magnetic permeability of multiphase materials. *J. Appl. Phys.* **33**, 3125–31

Hull D, Clyne T W 1996 *An Introduction to Composite Materials*, 2nd edn. Cambridge University Press, Cambridge, UK

Kelly A 1988 Advanced new materials; substitution via enhanced mechanical properties. In: *Advancing with Composites, Int. Conf. on Composite Materials*. CUEN, Naples, Italy, pp. 15–23

Kelly A 1994 *Concise Encyclopedia of Composite Materials,* revised edn. Pergamon, Oxford, UK

Kelly A 1996 The 1995 Bakerian lecture—composite materials. *Phil. Trans. R. Soc. London, Ser. A* **354**, 1841–74

Kelly A, Macmillan N H 1986 *Strong Solids*, 3rd edn. Clarendon, Oxford, UK

Kelly A, Rabotnov Y N 1983 (eds.) *Handbook of Composites,* 4 Vols. North Holland, Amsterdam

Kelly A, Clyne TW, Bonfield W, and Willoughby A (eds) 2006, *Phil. Trans. R. Soc. Lond.* A 364 (1838), Discussion Meeting Issue "Engineered Foams and Porous Materials": 3–281

Maire, E., Fazekas A, Salvo L, Dendievel R, Youssef S, Cloetens P, Letang JM 2003., X-ray Tomography Applied to the Characterization of Cellular Materials - Related Finite Element Modeling Problems. *Composites Sci. Technology*, **63**(16): 2431–2443.

Naslain R, Lamon J, Doumeingts D (eds.) 1993 High temperature ceramic matrix composites. *Proc. 6th European Conf. on Composite Materials*. Woodhead, Cambridge, UK

Parkhouse J G and Kelly A 1995, The Random Packing of Fibres in Three Dimensions *Proc. R. Soc. Lond.* A 451: 737–746.

Parkhouse J G and Kelly A 1998, The Regular Packing of Fibres in Three Dimensions *Proc. R. Soc. Lond.* A 454: 1889–1909.

Schüth F, Sing K S W, Weitkamp J (eds.) 2002 *Handbook of Porous Solids (5 volumes)* Wiley VCH, Weinheim.

van Suchtelen J 1972 Product properties: a new application of composite materials. *Philips Res. Rep.* **27**, 28–37

van Suchtelen J 1980 Non-structural applications of composite materials. *Ann. Chim. Fr.* **5**, 139–53

Talreja R, Månson J A E 2000 (eds.) Polymer matrix composites. *Comprehensive Composite Materials,* Vol. 2. Pergamon, Oxford, UK

Warren R 2000 Carbon/carbon, cement, and ceramic matrix composites. *Comprehensive Composite Materials,* Vol. 4. Pergamon, Oxford, UK

A. Kelly[a] and A. Mortensen[b]
[a] *University of Cambridge*
[b] *Ecole Polytechnique Fédérale de Lausanne*

This Page Intentionally Left Blank

A

Aircraft Materials

The selection of materials for aerospace is a complex decision-making process, always involving tradeoffs. It involves consideration of numerous factors such as environmental (temperature, corrosion) loading parameters, stiffness, durability, and damage tolerance. Cost, which includes raw materials and all subsequent processing costs, is becoming more dominant in today's global market, and the industry is directing an increasing portion of the materials development effort to this area.

1. Historical Development

Prior to World War II, aluminum was used almost exclusively in metal aircraft construction with some steel in selected areas where its higher strength and stiffness was required, such as landing gear and engine support structure components. With the development of the Kroll process as a production means of economically extracting titanium from its ore, titanium began being used for aerospace applications in the 1950s, driven primarily by the engine companies. It is now a key material for both airframe and engine structure due to its high specific strength and corrosion resistance relative to aluminum. In the early-to-mid 1960s, composite materials consisting of boron-based fibers impregnated with various polymers were introduced into aerospace structure. These composites soon evolved into the present-day carbon fiber composites, commonly referred to as CFRP (carbon fiber reinforced plastic) or GR/EP (graphite/epoxy) or PMC (polymer matrix composites). (The earlier boron fiber composites were very costly, and very difficult to machine. Development of the graphite fibers then displaced the earlier boron composites.) The polymeric matrices, graphite fibers, and processing have constantly evolved to the point where cost and structurally efficient airframe structure can be produced. They are attractive materials due to their low density, high strength and stiffness, and excellent fatigue characteristics in conjunction with the ability to tailor the composite layup to the specific requirements of the component. Thermoset epoxy-based resin systems typically have dominated the airframe industry. However, more specialized systems have been used for applications requiring higher operating temperatures, such as bismaleimides (BMI). Significant advances have also been made with regard to the development and application of aluminum (ARALL and GLARE) and titanium hybrid laminates (polymer matrix composite/Ti), and aluminum and titanium metal matrix composites, though their usage has been limited by high costs.

2. Temperature Considerations

Subsonic airframes are exposed to a wide range of temperatures; exterior structures exposed to the atmosphere can reach −54 °C to −73 °C while structures in the nacelle area may reach several hundred degrees. Typically, 7075-type high-strength aluminum alloys are used up to about 120 °C. Moderate-strength 2024-type alloys can be used up to 175 °C for short times, and the 2219 alloy can be used up to about 205 °C with the same caveat. High-strength low-alloy steels such as 4330M, 4340, and 4340M (300M) may be used up to about 205 °C. This limitation is governed by three factors: (i) restrictions due to the finish used for corrosion protection, (ii) retention of the residual stresses due to shot peening, and (iii) must be used at least 28 °C below the tempering temperature to prevent overtempering. Stainless and corrosion-resistant steels such as 15-5PH and PH13-8Mo may be used up to temperatures of $\sim$315–480 °C. Iron and nickel-based heat-resistant alloys are used at temperatures up to 705 °C and higher (most airframe structures experience temperatures below about 95 °C to 150 °C, with structures in the engine areas reaching about 370 °C.

Alpha or super-α alloys such as Ti–6Al–2Sn–4Zr–2Mo + Si are used at temperatures up to 540 °C for structural applications. The plug and nozzle on some of the newer transport engines use a titanium alloy (Beta-21S) at temperatures up to 705 °C for short times. The temperature is normally in the 540 °C to 565 °C, and lower, range.

Fighter aircraft operate over a similar temperature range as commercial aircraft. However, supersonic-capable aircraft have higher temperature requirements.

The existing 80 °C-cure epoxy PMC materials are presently characterized for use at temperatures to $\sim$35 °C. Higher operating temperatures have been demonstrated on military applications. They have pushed epoxies to 135 °C (over limited load-bearing structures) and BMI to a continuous operating temperatures of 165 °C, with higher excursions in isolated cases. The reader must note that military applications have lives that are typically less than half that of commercial aircraft, from a flight hour standpoint.

The processability of the epoxies and BMIs is typically well-known and characterized, and both systems are commercially available. Certain systems have a desired characteristic, such as controlled flow, for example. This allows for panels to be fabricated utilizing various low-weight core materials (honeycomb) to provide panel stiffness. Nomex or fiberglass core impregnated with phenolic resin is extensively used in structural applications. Controlled flow allows the use of this lightweight honeycomb without resin flowing into the cells, and subsequently resulting in a low fiber volume fraction in the facesheet plies. Alternate core materials are also utilized based upon the design/cost/performance requirements.

---

* Cross references marked by an asterisk are included in this volume.

[#] Cross references marked by a hash can be found by consulting the Encyclopedia of Materials: Science and Technology.

Specialized materials such a polyamides have been used in excess of the above stated temperatures. Several engine exit and compressor stator vane components have been fabricated. In addition, specialized structures that are "washed" with engine exhaust gases have also been produced.

3. *Design Considerations*

The design of aircraft structure must take into account four major failure modes

- Static
- Stiffness
- Durability, considering both fatigue and damage tolerance (crack growth, residual strength, and damage detection)
- Environmental resistance (corrosion and temperature considerations)

With the recent refinements of analysis methods, the safety factor of 1.5, component verification, and airplane proof and ultimate tests, there is an extremely low probability of a static failure of any undamaged structure. Many airframe components are designed by elastic stiffness or specific stiffness (modulus/density) considerations to minimize weight. With the widespread introduction of dissimilar composite and metallic structural materials, the requirement to match stiffness properties can be an important design criterion. Similar to the static analysis gains, technology improvements in design and analysis to meet the fatigue life requirements have minimized these concerns, though problems still exist in certain areas. Most of the fatigue problems in service have been on aircraft that have exceeded the design service objective (DSO). Safe operation of aircraft also hinges on design for damage tolerance. Commercial aircraft inspection periods are determined by the detectable defect size, crack growth rate and critical crack length. It must be demonstrated, by test or analysis, that if a crack of the minimum detectable size initiates between inspection periods, the aircraft may be operated until it is detected. Military aircraft take this concept one step further; defects are assumed present when the aircraft rolls out of the door and analysis must demonstrate that the aircraft can operate safely for two lifetimes (and actually test for four lifetimes) with this assumed flaw in the most critical location of a given component. This concept is designed to be safe with extensive damage—it *assures* the operator that they will have sufficient opportunity to detect and repair any damage before it becomes flight critical. A large portion of US Air Force aircraft structure is designed with this criterion. The US Navy tends more towards a safe-life approach, where aircraft life is designed by the number of cycles (flights) to initiate a detectable fatigue crack.

Critical requirements for a transport-type aircraft are summarized in Table 1.

The above criteria are specifically for transport-type aircraft, military or commercial. The differences in design considerations between military and commercial transports are relatively minor, the most significant being the need for rough runway capability for the former, and military transports are generally designed for a significantly shorter number of flight hours and the design is more driven by damage tolerance considerations. The differences are more substantial for fighter-type aircraft which also have shorter design lives (8000 hours as opposed to something of the order of 40 000 hours) and significantly higher loads, which means that design considerations are quite different, with damage-tolerance considerations driving US Air Force applications and fatigue-crack initiation driving US Navy designs.

Many steps have been taken to improve the corrosion protection of aluminum and steel alloys but this is still a maintenance issue with aircraft of today—though they are much improved in comparison to older aircraft.

Organic matrix composite materials are used to reduce weight, improving structural efficiency, and the design considerations for composite versus metallic structures are quite different. The primary difference in certifying these structures for service is accounting for barely visible impact damage for the composites. They can sustain significant damage that is not visible to the naked eye. The design allowable used to size structure is premised on several considerations, which includes barely visible damage, where applicable. It must be demonstrated that the airplane is capable of making a safe landing after the composite structure suffers an obvious in-flight discrete-source damage event. This would include events as severe as loss of a fractured engine turbine blade going through the structure. Potential environmental degradation must also be accounted for. Similar to the metallic structure, in addition to analytical verification, testing is carried out at the coupon, element, subcomponent, and component levels as well as full-scale limit-load testing to verify damage and environmental effects such as fluid sensitivity, moisture absorption, UV radiation, and oxidation.

4. *Materials Used*

Aluminum, steel, and titanium alloys and composites comprise ~99% of the structural materials used on

Table 1
Critical requirements for transport aircraft component design.

Wing	
Lower surface	
Skin (2XXX plate)[a]	Damage tolerance, fatigue
Stringer (2XXX extrusion)	Fatigue, damage tolerance, tension strength
Upper surface	
Skin (7XXX plate)	Compression strength (damage tolerance for heavy-gage skins)
Stringer (7XXX extrusion)	Compression strength
Ribs	
Shear tied (7XXX plate)	Shear strength
Intermediate (7XXX sheet)	Stiffness, shear strength
Fuselage	
Monocoque	
Skin (2XXX sheet)	Fatigue, damage tolerance, corrosion resistance
Stringer (7XXX sheet)	Fatigue, compression strength
Frames (7XXX sheet)	Stiffness, fatigue, compression strength, damage tolerance
Floors	
Beams (extrusion, sheet)	Static strength
Seat tracks	Corrosion resistance, static strength
Bulkheads	Fatigue, damage tolerance
Stabilizer	
Lower surface	
skin (7XXX plate)	Compression strength
stringer (7XXX extrusion)	Compression strength
Upper surface	
skin (2XXX) plate	Tension strength, damage tolerance, fatigue
stringer (2XXX extrusion)	Tension strength
Fin	
Skins (7XXX plate)	Compression strength, damage tolerance
Stringer (7XXX extrusion)	Compression strength

[a] 2XXX and 7XXX refer to 2000 and 7000 series aluminum-type alloys, respectively.

commercial aircraft. Advantages and disadvantages of these materials systems are indicated in Table 2. Some nickel-based alloys and specialty materials are used for specific applications.

The primary structural construction material for commercial aircraft, transports, and bombers (with the exception of the B-2) is aluminum alloys. (The wing/fuselage skins of the B-2 are PMC, using an unmodified 175 °C-curing epoxy resin matrix.) The materials breakdown for transport/cargo type aircraft is shown in Fig. 1; it can be seen that the trend is towards increased composites and titanium usage, primarily at the expense of aluminum. It should be noted that 2.5% of the 78% aluminum on the C-17 is Al–Li (2090 sheet and extrusions used in the cargo floor area). The materials selection is much different from that of fighter aircraft (Fig. 2). As can be seen, the materials mix varies widely for fighters depending on their missions. Current day fighters tend to have composite skins, operating temperature permitting. A significant change in the F-18 materials mix occurred as the performance requirements were increased. Titanium and PMC usage increased substantially at the expense of aluminum structure.

Aluminum alloys have been the backbone of aircraft construction since just prior to 1920, and most of the advanced alloys are simply variants of 2024, which was introduced in 1921, and 7075 which was introduced in 1943. Refinements in the chemistry, processing, and heat treatment have led to the evolution of properties illustrated in Fig. 3. There has been a continuous evolution of 2024 with emphasis on improving fracture toughness, fatigue-crack initiation and crack-growth resistance; with 7075 type alloys the emphasis has been to improve the strength (primarily compression strength) and fracture toughness, without sacrificing the other properties in either case. Along with the cited desired performance improvements, means of improving corrosion resistance are always of interest. The heat treatments selected for both alloy types have sacrificed some strength capability for the sake of corrosion resistance.

Table 2
Advantages and disadvantages of specific alloy systems.

Advantages	Disadvantages
Aluminum alloys	
Low density; good mechanical properties; low cost; low manufacturing costs; good toughness, $\mathrm{d}a/\mathrm{d}N$, fatigue properties combinations; excellent machining characteristics	Corrosion; high-strength alloys not weldable[a]; high-strength alloys only used to 95 °C; not compatible with composites[b]
Titanium alloys	
High strength/density ratio; corrosion resistance; high operating temperatures; compatibility with composites; SPF; SPF/DB capability	High initial cost; high machining costs; difficult to form; high notch sensitivity
Low-alloy steels	
Highest strength; moderate toughness; low raw material cost	High density; corrosion; rigorous heat treatment and machining requirements for highest strengths
Corrosion-resistant steels	
Corrosion resistance; moderate-to-high strength; good fracture toughness; reduced manufacturing costs/flow time relative to low-alloy steel.	High density; low strength relative to the high-strength low-alloy steels
Nickel and heat-resistant alloys	
Moderate to high strength; excellent corrosion resistance; high operating temperature	High density; high cost; difficult to form and machine
Polymer matrix composites (PMC)	
Light weight; low density; high stiffness; high strength; can "tailor" properties; fatigue not an issue	Property characterization due to moisture equilibration and UV absorption; possibility of significant undetected damage; galvanic incompatibility with Al and low alloy steels

[a] Most aluminum–lithium alloys are weldable. [b] Corrosion protection must be used to fabricate Al/PMC structure.

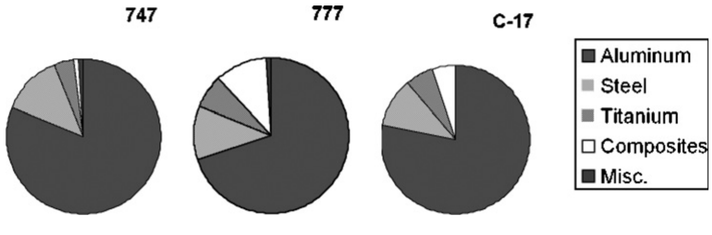


Figure 1
Structural materials usage: transport-type aircraft.

The latest alloys utilized on the Boeing 777 include:

> Upper wing—skin, 7055-T7751, high strength, good corrosion resistance; extrusions 7055-T77511.
> Lower wing—skin, 2324-T39; stringers, 2224-T3511.
> Fuselage—skin, 2524 (still 2024-T3 on other Boeing commercial aircraft); upper crown stringers, 7150-T77511; lower surface stringers, 7075-T6. The primary forging alloys are 7075-T73 and 7050-T74. The latter alloy is higher strength than the former. It also provides better properties in heavy sections. The high-strength aluminum alloy used on the C-17 was primarily 7150 in the T7751 temper. The MD-11 utilized the same alloy in the higher strength T6 condition.

Alloy 6013 is used for formed parts of secondary structures in place of 2024 due to its superior formability (the fatigue performance is slightly inferior to that of 2024). The 6013 alloy can be formed in the T4 condition and aged to the T6, whereas 2024 must be

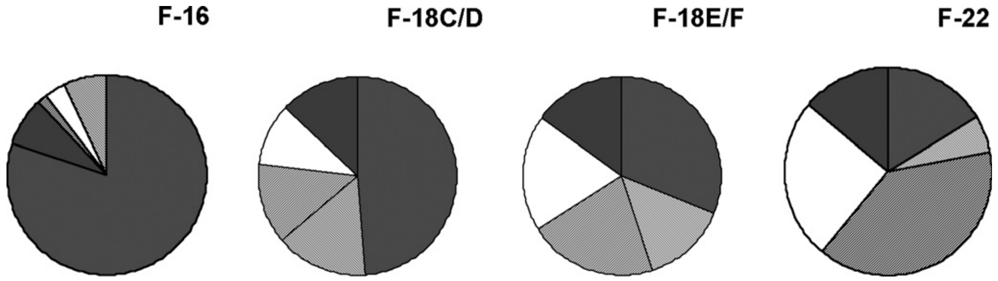


Figure 2
Structural materials usage: fighter aircraft.

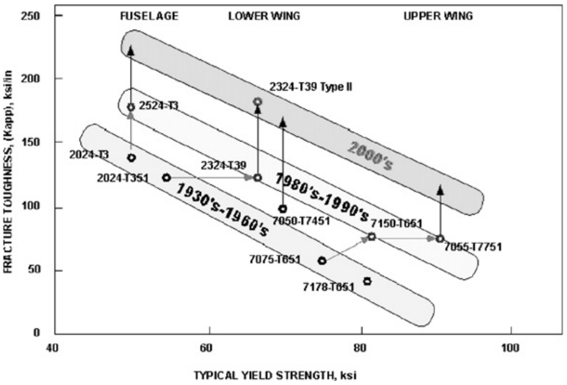


Figure 3
Aluminum alloy evolution.

formed in the "O" condition for optimum formability, then solution treated and aged. The requirement for solution treatment often creates distortion problems with formed parts. The 2024 and 7075 alloys can be procured in the "Hi-form" condition, a processing route developed by Kaiser, to provide a finer grain size and better formability. These can be formed in the "W" condition, to bend radii comparable to the conventional alloy in the "O" condition and aged to the proper temper. If formed in the "O" condition, it can be formed to tighter bend radii than that of the conventional material in the "O" condition, but it must be solution treated and quenched prior to aging, creating the potential for distortion.

The aluminum alloys used on fighter aircraft are similar to those on commercial aircraft, primarily 7050, 7175, and 2124. Major bulkheads typically use thick plate, such as the 2124-T81851 used on the F-16 and F-22. Additionally, 7000 series alloys 7075-T76 and 7475-T761 and 2000-series alloys 2024-T81 and 2124-T81 are used for sheet applications, while 2124-T81 plate is used for plate applications. Additionally, 6013 sheet has potential for future applications because of its improved formability (as defined earlier), excellent exfoliation corrosion resistance, and lower cost.

The first Al–Li wingskin (8090 T851) was flown on an F-15 as a demonstrator. Al–Li has not been used since then in production for these types of applications. Thick-section 2097 Al–Li alloys offer significant benefits for aircraft with severe fatigue requirements due to reduced cracking rates and alloy density benefits.

The 5XXX series Al–Mg–Sc alloys exhibit similar mechanical properties to 2024-type aluminum alloys, while perhaps avoiding the need for protective alclad surface finishes.

Historically, the prevalent structural steel alloys used on commercial aircraft are the high-strength low-alloy steels such as 4340, 4330M, and 4340M. These alloys have been used in situations requiring very high strength (as high as 1900–2070 MPa), such as the landing gear and flap tracks. While these materials are structurally efficient, they require extra attention to design and protective finishes to address general corrosion and stress-corrosion issues. Despite protective finishes such as Ti–Cd plating, primer, and enamel, in-service inspection and maintenance are still significant.

More recent commercial models and derivatives have increased the use of high-strength corrosion-resistant steels such as 15-5PH wherever possible. Relative to the high-strength low-alloy steels, corrosion-resistant alloys like 15-5PH reduce processing flow time and cost in production (the processing of high-strength low-alloy steels, particularly machining, must be accomplished within a narrow window and is quite time consuming and costly). The need for protective finishes is also minimized which decreases cost, flow time, and environmental impact (the raw material costs are on a par with the low-alloy steels.) The primary limitation of these alloys has been strength (1240–1380 MPa is the maximum ultimate tensile strength for 15-5PH). Development of new corrosion-resistant alloys which come closer to matching the strength of high-strength-low alloy steels will increase the use of these alloys.

Due to stress-corrosion and fracture-toughness requirements, imposed primarily by the US Navy, the steel alloys used for military applications tend to be a little more "exotic." Primary landing gear components are fabricated from Aermet 100, which is used at a tensile strength level of 1930–2070 MPa. Preference for this alloy is driven primarily by its high toughness (minimum K_{Ic} requirement of 110 MPam$^{1/2}$) and improved stress-corrosion resistance. (4340M has a typical K_{Ic} of ~65 MPam$^{1/2}$.) The Aermet alloy is also weldable, and can be used up to temperatures of about 425 °C. Its attributes would be attractive in the commercial sector, but the alloy is considerably more costly than 4340M. Reduced maintenance could result in lower life-cycle costs, but these can be difficult to assess. For the lower-strength (1415–1585 MPa) steel damage-tolerance applications, PH13-8Mo is utilized. This alloy has higher strength and toughness than 15-5PH, with superior thick-section properties. Some 15-5PH is also used for noncritical applications—it has lower cost than PH13-8Mo.

Titanium alloys used cover the entire range of conventional titanium alloys commercially produced. From commercially pure titanium, where formability and corrosion resistance drive its use, to α- or super-α-type alloys such as Ti–6Al–2Sn–4Zr–2Mo for high-temperature applications, to α/β alloys such as Ti–6Al–4V, which is the workhorse alloy of the aerospace industry, and β alloys such as Ti–10V–2Fe–3Al. Ti–6Al–4V was developed in the late 1950s and still represent ~80% of the aerospace market. It is a moderate-strength alloy (with a minimum UTS of about 895 MPa) which is quite forgiving in its processing, exhibits good properties combinations and can be utilized for every product form. The properties of the alloy can be tailored to the application. For instance, a β-annealed extra-low-interstitial (ELI) grade is used for damage-tolerance applications—this material is used extensively on the F-22. With a slight drop in tensile properties relative to the annealed standard grade material, the fracture toughness can be doubled and the fatigue crack growth rate decreased by an order of magnitude—however, this is achieved with a significant decrease in fatigue crack initiation life.

The β-alloys can generally be heat treated to higher strength, have deeper hardenability, and are generally easier to fabricate into a semiproduct form. These alloys tend to be utilized for specific product forms. Ti-10V–2Fe–3Al is a high-strength forging alloy, used typically at a minimum tensile strength of 1170 MPa and somewhat higher for gages less than 25 mm. (There has been a reluctance to use alloys such as this in Europe due to potential part distortion during machining because of residual stresses from the water quench associated with the solution treatment.) Ti-15V–3Cr–3Al–3Sn is used primarily for sheet (strip)—as a β-type alloy it is strip producible, which is a significant economic advantage, particularly for thin gages. This product is used at a minimum tensile strength of 1035 MPa and higher. Titanium is an excellent spring material—with a density ~40% lower and shear modulus about half that of steel, a titanium spring can provide weight savings of up to 70% and volume reduction of up to 50% in comparison to steel, and eliminate the corrosion problems prevalent with most steel springs. Ti–3Al–8V–6Zr–4Cr–4Mo (β-C) is used for springs at a minimum tensile strength of 1240 MPa (flat springs are fabricated from Ti–15V–3Cr–3Al–3Sn). β-21S is a high-temperature alloy used in sheet and foil form for exhaust nozzles. As mentioned, these structures see temperatures as high as ~705 °C for short durations—the temperature is below 540–565 °C for normal operation.

The titanium alloys used are similar for commercial aircraft and military transports/bombers and fighters, but the heat treatments may be different due to the emphasis on damage tolerance. Ti–6Al–4V is the primary titanium alloy used for damage-tolerance (DT) applications. The ELI grade is used for these types of applications in the recrystallize annealed (RA) or beta annealed (BA) condition. The BA condition has the best DT performance as mentioned previously for commercial aircraft, but poor fatigue crack initiation resistance. This heat treatment is used extensively on the F-22. The fracture toughness and fatigue crack growth resistance is not quite as good in the RA condition but the fatigue behavior is better. This heat treatment is used on the B-1, B-2, and F-18.

Ti–6Al–6V–2Sn is used on some of the older fighter aircraft. It is generally used in the annealed condition, giving it a tensile about 70 MPa higher than that of Ti–6Al–4V, which provides some weight savings. It is not used on newer aircraft with the increased emphasis on DT, as this alloy has lower fracture toughness and stress-corrosion resistance than Ti–6Al–4V.

The processing for another, somewhat higher-strength alloy, was developed for use on the F-22, Ti–6Al–2Sn–2Zr–2Mo–2Cr. It is used at a minimum tensile strength of 1035 MPa and fracture toughness of 77 MPam$^{1/2}$ with significantly superior fatigue performance, which, again, provides some weight savings over Ti–6Al–4V. It is used for the lower keel cord and intermediate wing spars on the F-22. Greater usage is anticipated as experience with the alloy is gained. Ti–6Al–2Sn–2Zr–2Mo–2Cr is the primary structural titanium alloy (everything except sheet) for the X-33, NASA's half-scale demonstrator reusable launch vehicle.

The F-22 makes extensive use of welding for fabrication of primary structure (the forward and aft booms).

Titanium is an attractive material for aerospace applications, but its cost is quite high. Raw material costs may be several times higher than aluminum or steel, and machining times are of the order of 10 times as great in comparison to aluminum. This precludes using this material except where the added cost can be justified through improved structural performance. (This is one of the drivers for greater utilization of castings, as parts closer to net shape, incorporating multiple parts with less machining and assembly time, can significantly reduce the cost of titanium hardware.)

In general, polymer matrix composites are used for reduced weight. They have strong advantages with regard to fatigue performance and corrosion resistance. Another key advantage is that the layup can be tailored to provide the desired properties in the critical locations/orientations.

Historically, the utilization of composites has been somewhat restricted because of its high cost. However, the situation is now changing. A good design, and taking advantage of the processing benefits of PMCs over metallic structure, can result in lower cost structure. With the development of automated material placement, the use of computers to define the specific ply layups, advances in the resin transfer molding process, etc., the costs of organic matrix composites is coming down, and cost savings are being reported. A good example is the C-17 empennage. It has 2000 fewer parts, 42000 fewer fasteners, and is 215 kg lighter. These advantages, and the requirement for 1000 fewer tools, resulted in a cost saving of ~40% in comparison to a conventional built-up aluminum empennage. British Aerospace has indicated that processing advances along with the computer aided design of the layups enabled them to cost-effectively utilize PMC composites for the front and rear fuselage and wing assemblies of the Eurofighter Typhoon.

As mentioned previously, the PMC materials used on aircraft usually entail carbon fibers in a broad range of polymer matrices (epoxies, PEEK, BMI, phenolics, etc.). (The F-15 is a bit unusual in that the bulk of the 1.4% composite material used on that aircraft was the older technology boron fibers in an epoxy matrix, which was used in the vertical fin and horizontal stabilizer.) There is also a significant amount of glass-fiber reinforced epoxy composites for secondary structures on commercial aircraft.

The thermoset PMC structure on the F-22 consists of about half BMI and half epoxy resin parts, both tape and fabric forms, with IM7, Nicalon, and S-2 glass fibers. Toughened epoxy PMCs and BMI and epoxy RTMs are also utilized. Thermoplastics are being used for specialty items such as landing gear and weapons bay doors. A drawback to these materials (capable of higher-temperature service) has been that they typically exhibit increased difficulty in processing relative to the lower-temperature PMCs. Higher process temperatures, Kapton bagging films, low coefficients of thermal expansion, CTE, and tooling have required that modified consolidation procedures are utilized. These solvated resin systems polymerize by a condensation reaction mechanism. Processing requires high-temperature cure cycles that can vaporize the low-molecular weight solvents and reaction products, and maintain resin viscosity low enough to achieve adequate resin flow for complete saturation of the part (100% density).

The Lockheed/Boeing F-22 is the first aircraft, commercial or military, to take advantage of resin transfer molding for fabrication of primary structure composite parts. This process is used for over 150 parts, ranging from the lip edges of the engine inlet to sine wave wing spars.

5. *Future Trends*

5.1 Aluminum

Development efforts in the near future will focus on several key areas. In terms of increased performance, efforts will be directed toward wing alloys with better fracture toughness at equivalent strength levels, and Al–Li alloys for weight savings. Airbus will use GLARE, an aluminum/fiberglass PMC with good fatigue, corrosion, impact, and fire resistance for fuselage structure on the A380. The size of this large airplane, and its associated weight, probably dictates implementation of new, higher-performance materials.

Cost-reduction efforts will include development of alloys which can be formed in the tempered condition and heavy-section alloys with more isotropic properties for monolithic structure. There is increasing interest in alloys and processes that integrate parts to a larger, more monolithic structure, such as castings. These provide the possibility of reducing cost, and sometimes weight, through reduced machining, and reduced part, tool, and fastener counts.

Increased use of welding in areas where it has not been possible before also address reducing the number of fasteners and the labor associated with their installation; hundreds of thousands of fasteners are used on each fuselage. Friction-stir and laser welding are being considered for aluminum fuselage sections. Airbus will use stringers laser welded to skins on the lower shell of the fuselage of the A318 and A380 (a new large airplane capable of carrying 555, or more, passengers). Their goals for welded structure are to reduce the cost by 25% and the weight by 15%. Friction-stir welding is also a high priority at Airbus.

5.2 Titanium

Improved-performance titanium alloy development will focus on alloys suitable for providing high strength combined with increased producibility. Titanium matrix composites (TMC) provide attractive properties, particularly for stiffness-driven applications, but cost is still an issue. The first TMC application, titanium actuator pistons, has been implemented on the Pratt and Whitney engine for the F-22, and General Electric has a titanium matrix component for the F-16 engine. The F-22 program is working with the supplier to determine the feasibility of utilizing this technology for stiffness-critical airframe structure. Titanium hybrid laminates are also being studied, but their development is still embryonic in nature.

However, the bulk of the development efforts for titanium will be directed toward reduced cost. Permanent mold castings, both gravity metal and vacuum die-casting processes, show substantial promise for reducing costs for smaller, less complex castings. A new powder metallurgy process is being developed for producing relatively simple shapes via a laser process, which utilizes a stereolithography-type technique to produce a fully dense metal shape. This process is being studied for the production of machining and forging preforms.

The possibilities of further utilizing SPF and SPF/DB for fabricating complex structures are being analyzed as another means for saving cost.

5.3 Steel

Development efforts in the steel area are focusing on two primary areas; (i) improving the damage tolerance attributes of high-strength CRES at current and higher strength levels, and (ii) increasing the strength of the corrosion-resistant alloys to replace current high-strength low-alloy steels (e.g., 4330M, 4340M/330M). A current example is Carpenter's Custom 465 which is a high-strength CRES with potential for use up to 1585–1725 MPa. Work is ongoing to develop other CRES alloys with strengths approaching that of 4340M/300M, i.e., 18995–2070 MPa ultimate tensile strength.

5.4 Organic Composites

The primary development efforts in the organic composite arena will also focus on cost, via raw material and processing improvements. Focuses to accomplish this include reduced cost materials (achieved via increased volumes and improved compositions), improved assembly methods such as co-cure, thermoplastic welding, and enhanced processing/fabrication procedures. Technologies under investigation here include greater utilization of automation, larger tow placements, nonautoclave cure (E-beam), and adaptive tooling.

Technologies being studied/implemented providing both cost and performance benefits include noncrimp fabrics, stitching, and near-net-shape preforming. Processing routes to provide these benefits include resin transfer molding (RTM/closed mold and VaRTM/vacuum bag), stitched resin film infusion, liquid resin infusion, and resin transfer molding. Considerable effort is also being devoted to improving the environmental durability, to provide better UV resistance, durable paints, repair, etc.

R. Boyer
Boeing Commercial Airplane Company, Seattle, Washington, USA

Alternative Techniques for Manufacturing Composite Aerospace Designs, Cost of

Aerospace companies are in a continuing battle to increase the performance and reduce the weight of their aircraft structures, while lowering acquisition and operating costs. Thus, there is always interest in new materials with higher performance and lower manufacturing and ownership cost than aluminum, which is widely used in aerospace manufacture because of its strength and weight. A new material that offers to deliver these benefits is composite materials. Composite materials are made up of two components, a fiber or mat of fibers that offer stiffness, strength, and shape to the structure, and a binding material that helps to maintain the integrity of the fibers. Usually, the binding material is responsible for few of the mechanical properties of the composite system.

Although there are many possible fiber and resin material combinations, for current aerospace structures the fiber material used is almost exclusively carbon (graphite). Epoxy resins tends to be used as the binder. Other systems may be used once technical, cost, and certification issues have been overcome. It should be noted, however, that engine makers are also interested in composite materials, both carbon/epoxy systems for structures like nacelles and more advanced fiber/binder combinations (such as ceramic matrix composites) for high-temperature applications like turbine blades.

Since the introduction of the carbon fiber/epoxy composite system, each new generation of aircraft has seen a slow but steady increase in the use of composites. Nowadays, military aircraft use considerable amounts of composite materials (over 30% of the structural weight), as do light aircraft and business aircraft. Indeed, an all-composite business jet has been produced. Civil airliners have traditionally used less composite materials, where to date it has mainly been used in secondary structures such as slats and doors. The main hindrance to realizing the performance and weight advantages of composite materials has been cost; despite the promise and possibility of low-cost manufacture of composite components they have stubbornly remained more expensive than their aluminum counterparts. Reducing the cost of composite structures has been the focus of major research programs in the USA, the EU, and the Far East.

Single carbon fibers are never processed individually; the carbon fiber manufacturers provide them as tows, tapes, or fabrics. Tows are long bundles of carbon fibers, with currently 8000–24000 fibers in a tow. A greater number of fibers in a tow make it cheaper but there is less control over the number of fibers, so aerospace designers tend not to use the tows with the highest number of fibers until the manufacturers can ensure a consistent performance. As the name suggests, tapes consist of long, unidirectional bundles of fibers, currently up to 150 mm in width for aerospace applications. Tows and tapes are usually provided on an adhesive backing paper, which helps to maintain the shape of the tows and tapes during transit and storage. However, the backing paper adds an extra cost and the tow placement and tapelaying processes need to introduce steps to remove the backing paper prior to placement, which adds complexity and cost.

Fabrics are tows woven into a sheet, like the fabric of clothing. The fabrics can be crimped, so that the fiber layers are interwoven like a plaid or they can be laid in single layers, with the shape maintained by stitching the layers together. Fabrics can be over 1 m in width. The fiber layers are oriented at varying angles; 0°, ±45°, and 90°. By altering the number and order of layers of various orientations, the designer can design the composite material to optimally manage the stresses and loads in the structure.

Epoxy resins are generally used as the binding agent because they offer an attractive combination of strength, ease of processing, and cost. The resin can be applied to the fibers in a variety of ways. The manufacturer can coat the carbon fiber with resin (pre-preg) or the resin can be added to the structure as part of the manufacturing process. The resin needs to be heated (or cured), usually, in a pressurized vessel (an autoclave) for the resin to acquire its mechanical strength. The autoclave adds to the cost of the process, but equally significantly it can constrain the size of parts that can be processed and the factory logistics and manufacturing lead times are affected by the need for most of production to use the autoclaves. The curing process introduces some risks to the manufacturing process; uneven heating of the structure and the tooling in the autoclave may distort the final shape of the part. As a consequence, there is much interest and research in resin systems that can cure at lower temperatures, possibly even at room temperature.

Composite material is considerably more expensive than the equivalent weight of aluminum. The cost of carbon fiber varies depending upon how it is provided but typical values would be around \$100 kg^{-1} for resin impregnated fabrics and \$90 kg^{-1} for tape, although high quality or low volume products will cost more. In contrast, aerospace grade aluminum typically costs around \$20 kg^{-1} (again, the costs depends heavily on the quantity and precise grade). Consequently, for composite structures to be cost competitive they must be lighter (hence needing less material) and more cost-effective to manufacture. The superior properties of composites mean that they can usually perform the same structural or aerodynamic function with less material, especially as aerospace designers learn to fully exploit the properties of the material. With weight an important determinant of fuel consumption, less weight can make a major contribution to ownership costs and help to justify the use of composite structures. The manufacturing technologies also need to ensure that the cost savings in the process more than compensate for the higher material cost. They must also address material waste; because composite material is more expensive than aluminum, and currently cannot be recycled, all the manufacturing processes must focus on waste material and reduce it to a minimum.

As carbon fiber and epoxy resin technologies are mature processes with their major customers outside aerospace, the major cost savings will be from new materials, new manufacturing processes, or (most likely) a combination of the two. That has been the case to date, with new manufacturing processes prompting new ideas for fiber and resin preparation and vice versa.

One of the earliest manufacturing processes used in aerospace manufacture, and still one of the most common, is manual or hand layup. The process begins with a fabric that is cut to shape on a cutting table, usually using a numerically controlled (NC) ultrasonic knife or water jet. To achieve the required thickness of material and the shape of the structure, multiple layers of fabric are usually necessary. The fabrics are positioned in a mold-tool, whose inside surface mirrors the required outside surface of the structure. Once the fabrics have been built up to the required shape, an inflatable bag is positioned over the structure and the resin is cured in an autoclave. The purpose of the bag is to apply pressure to the structure, which helps to maintain the shape and to drive air and water out of the resin. After curing, the structure is unbagged and removed from the tool.

Another manufacturing process currently in common use in aerospace is tapelaying. This uses a CNC-controlled robot that accurately positions the tape on a mold-tool, creating the shape by building up the layers of tape. As the buildup is CNC controlled and uses a contoured head and thin tape, complex shapes can be produced. With many layers of tape, it is common to periodically remove the tool and compress the material (debulking), as the head (which needs to be light to accelerate quickly) may not provide sufficient pressure to compact the layers. Afterwards, the structure is cured and processed as before.

Rather than using tapes, filaments or tows of composite fiber can be used. The filament-winding process wraps a tow of carbon fiber around a rotating tool, with heat applied at the point of contact to soften the resin. This process can be used for cylindrical structures, such as fuselage panels.

Several newer composite manufacturing processes are been developed and are beginning to be used in production, i.e., resin film infusion (RFI) and resin transfer molding (RTM). RFI is similar to manual layup in that the fabric is first cut to shape by a cutter. However, unlike manual layup the fabric has not been coated with resin (only a thin layer to help the resin film bind to the fibers). The layers of fabric are laid in the mold-tool and interleaved with thin films of resin, possibly using a robot or similar automated system. When the structure has been built up, the structure is bagged and processed in an autoclave.

RTM again uses fabric cut to shape and placed in a mold-tool. However, the mold-tool is part of a vacuum system, and when the structure has been built up the mold-tool is sealed, evacuated and heated. Resin is pumped into the structure and the combination of heat and vacuum enable the resin to flow through the structure. Unlike the other processes described, the structure does not need to be cured after processing, as the cure occurs *in situ*.

Pultrusion is widely used for the manufacture of low cost, constant cross-section products like pipes and gutters. Some aerospace structures, particularly stringers (and possibly skins) maintain nearly continuous cross-section and pultrusion processes to manufacture these components are under development. The precise details of the pultrusion processes vary, but they all pull fibers through a mold of constant (or nearly constant) cross-section and then through an in-line oven to cure the resin.

As discussed, each method varies in technical performance as well as cost, but the decision on which process to use should not be based solely on cost. Illustrative costs are shown in Table 1. Consideration of the respective cost of these processes is hampered by the fact that much of the data is held by companies and is confidential. However, enough experience in research and production has been gained for some conclusions to be drawn.

Table 1
Comparative costs of composite materials and processes.

Material/process options	£kg^{-1}
Unidirectional tape layup	129
Woven pre-impregnated tape laid by hand	104
Woven fabric hand laid by hand followed by RTM	72
Tow placement followed by RTM	50
Tow placement followed by RFI	61
Unwoven fabric machine laid using a laminator followed by RTM	59
Unwoven semi-impregnated fabric laid using a laminator and autoclave cured (similar to RFI)	105

The costs of the manual layup process are dominated by direct labor and material costs; little automation is used so the capital costs are proportionately low. This process tends to be most cost-effective for low-volume applications, where the production run is insufficient to absorb the cost of automated equipment. "Low volume" will depend upon the application, but for production runs greater than a few hundred, the long layup times and consequently high labor costs mean that manual layup tends not to be cost competitive with other manufacturing processes.

Unlike manual layup, the labor content of the tapelaying process is small, as the tapelayer needs little manning when in operation (some labor is needed to load the tape and to respond to breaks in the tape). However, the capital costs are high since the tapelaying head is a complex structure and the supporting structure needs to be quite large and rigid to provide dimensional control. Hence, tapelayers cost a few million dollars each, and if used for low volume production the cost of the process per part can be high (over twice the cost of competing processes not so reliant on expensive capital equipment). The tapelayer is advantageous if high volume production of parts is required as the volume can absorb the capital charges. Tapelaying can produce complex shaped parts and, because of the control of fiber orientation tapelaying makes possible, this process can produce the highest quality structures. Another advantage this approach offers is that it minimizes material waste.

Filament winding requires little in the way of capital or labor, making it in normal circumstances the cheapest composite manufacturing process. However, the restricted shapes that can be produced means that it has so far found few applications beyond making fuselage panels.

The RFI process was designed to achieve high deposition rates of composite material, a drawback of processes requiring tapes and tows; current RFI technology may achieve deposition rates of 100 kgh^{-1}, as opposed to 4–10 kgh^{-1} often seen for tapelaying. The use of fabrics, however, means that RFI has difficulty in making small complex shapes and its main application is likely to be in large, simple structures such as wing skins and spars.

Like filament winding, RTM needs little equipment and labor, the main equipment cost being the resin injection system and the tooling. As the tool needs to include a heating element and support an airtight seal, an RTM tool costs up to twice as much as conventional tooling. However, the autoclave curing is removed which makes this a very cost-competitive process. Most applications of RTM to date have focused on fairly complex parts. However, with further development and production experience, the combination of low cost and part variety that can be manufactured make RTM likely to become one of the major composite production processes.

These processes are not exclusive, and in some applications a "mix and match" approach would be the most cost-effective. Wing skins build up using RFI with hand layup for the small features such as pylon supports could be one example where the technology is selected for the application for which it is most suited.

Table 1 gives an indication of the relative costs of several composite manufacturing processes (based on research undertaken in the UK). The material types include tapes, tows, and fabrics (unwoven fabrics have planes of fabric held together by stitches, whereas woven fabric interleave the layers in order to maintain the shape). The manufacturing processes are tape layup, RTM, RFI, and hand layup. The laminator is used to transport the cut fabrics into the tool. The structure used as an example for this costing exercise was an aircraft tail skin.

Much of the potential savings from composite technologies lie in the possibility of major reductions in assembly costs. For metallic structures, assembly techniques have used highly specialized, highly expensive fixtures and jigs with parts assembled using mechanical fasteners (such as rivets). As well as the high capital outlay for the fixtures and jigs, the assembly process is very slow and labor intensive with long lead times and high inventory levels. Composite structures may offer two advantages. The part count can be significantly reduced, and bonding the structure using resins or adhesives can significantly reduce the assembly time and cost.

The part count can be reduced in two main ways. Complex shapes can be manufactured, especially with tapelaying and RTM, which combine in one structure several pieces that would have been made separately and joined if made of metal. In addition, processes such as RFI can make very large single piece structures (such as wing skins) which can replace numerous metal plates bolted together.

The second potentially major cost advantage of composite structures is that they can eliminate the mechanical fasteners needed to assemble the structure. Assembly can account for up to half the recurring cost of an aluminum structure, and much of this cost is in fastening. The structures can be bonded using adhesive and specially designed joints. Indeed, some of these assembly steps can be combined with part curing, so that parts of the structure are cured and bonded in the same operation. Eliminating 20–40% of the fasteners by using adhesive bonding offers a considerable cost saving.

However, for large civil aircraft many of these potential benefits have yet to be seen, and assembling composite structures has still tended to be more expensive than assembling similar metallic structures. The certification and testing of bonded structures requires a major effort to demonstrate the safety and reliability of the technique, especially for civil airliners, and until this is undertaken composite structures will still be mainly bolted. Composite structures are more expensive to drill and bolt, as special drill bits and bolts are needed, and there is the potential of damaging the structure through delamination. The composite structures sometimes require shimming, a process where the gaps caused by variances in the parts are filled either with inserts for large gaps or an adhesive if the gap is small. As this is a manual process, it can add considerably to the cost.

The tooling of composite parts differs from metallic parts in a number of respects. For most composite manufacturing processes, the resin is cured while the part is in the tool (as it is needed to help maintain the shape). As the tool and the structure are heated, the designer must take account of mismatches in the thermal expansion coefficients that may introduce stresses into the composite structure resulting in its shape being distorted when it is remove from the tool (springback). Composite tooling can be used to eliminate this problem, but composite tools have significantly shorter lives than their metallic counterparts, although their initial cost is lower. Uncured composite structures are sticky and the tooling for composite structures needs to be cleaned at regular intervals.

Aircraft with composites have the potential to offer significant improvements in ownership costs. In addition to their lighter weight, experience with composite structures shows that they demonstrate little material fatigue and reduced corrosion. Water ingress and delamination are concerns with composite structures, but with proper preparation and maintenance practices the performance of these structures does not diminish with time to the same extent as their metallic counterparts. Repairing composite structures is currently more expensive; costly autoclaves, inventories of different carbon materials and epoxies, and other significant infrastructure costs need to be provided. In addition, faults in composite structures such as internal delamination are difficult to inspect and repair. It is expected, nonetheless, that the cost differential between metallic and composite repair will narrow as users gain more experience in their use.

It can be seen that comparing the cost of composite manufacturing technologies is not straightforward; most manufacturing technologies have a niche in which they are likely to be the most cost-effective technology. Technologies using fabrics, such as RFI, are most suited for large, thick structures without complex shapes, whereas RTM may be more suited to smaller, more complex structures. Thus, the manufacture of composite structures on aircraft is likely to use several process technologies, suited to the applications where they have a cost (and technical advantage). For assembly, composite structures offer a significant cost advantage when structures are bonded, and the constraints that exist with bonded structures arise from testing and certification issues, as opposed to production concerns.

To conclude, carbon fiber composite materials have the potential to reduce the manufacturing and operating cost of aircraft, but challenges remain in developing the manufacturing technologies to realize this potential and in designing structures that fully exploit the potential of both the material and the manufacturing processes under development. So far, composite materials are primarily used on aircraft because of the performance and weight advantages it offers. It should soon be feasible for composite structures to compete on cost as well.

A. Williamson
Cranfield University, UK

Asbestos

Asbestos or amiante (from Greek *Asbestos* (inextinguishable) and *amiantos lithos* (unstained stone), respectively) is a commercial ancient name for fibrous, naturally occurring inorganic materials that are valued in the industry for their resistance to heat and high tensile strength (Michaels and Chissick 1979, Hodgson 1986, Habashi 1988, Virta 1999). The fibers are formed under rare conditions when certain silicate minerals crystallize in bundles of hundreds of thousands of strong, flexible fibrils that look like vegetable fibers. Chemically, they have the same composition as the other minerals found in the rock within which they are formed. Asbestos is an extremely useful natural resource, and has been studied extensively mineralogically and for industrial application. Its extraction from ores has been greatly improved with respect to safety in the work place.

Asbestos is usually present to the extent of about 5% in its deposits. To recover 1 ton (t) of asbestos, about 20 t of waste rock must be removed. This is achieved by fracturing the rock in crushers and other impact hammer mills to liberate the fibers. These are then fed to shaking screens equipped with suction hoods. Large fans provide the suction that lifts the fiber from the screen and carries it to collecting cyclones. A plant in Australia is said to be separating asbestos fibers from the rock by wet methods (Stewart 1986). Such a process would be interesting because it minimizes the presence of fibers in the work place. World production of asbestos is about 4 million tons per year, 55% of which is contributed by Russia and 20% by Canada. Each of the following countries contributes less than 4%: South Africa, China, Zimbabwe, Italy, Brazil, and the USA.

1. *Properties*

Asbestos has a relatively large specific surface area of 3–9 $m^2\,g^{-1}$. When suspended in distilled water (5 $g\,l^{-1}$) the pH sets at about 10 and Mg^{2+} can be identified in solution in an amount near to the solubility product of magnesium hydroxide. In the presence of acids, more magnesium can be solubilized and complete extraction can be achieved by long contact time or by boiling under a reflux condenser for about 2 h. Magnesium-free asbestos still retains its original shape but becomes fragile.

When heated in the temperature range 600–780 °C, chrysotile asbestos loses water due to dehydroxylation (an endothermic reaction). In the temperature range 800–850 °C, the dehydroxylated product undergoes crystallization to forsterite (an exothermic reaction). The overall reaction is:

$$2\ Mg_3(Si_2O_5)\ (OH)_4 \rightarrow 3\ Mg_2SiO_4 + SiO_2 + 4\ H_2O$$

The forsterite persists together with silica up to 1000 °C, but above this temperature some enstatite is formed:

$$Mg_2SiO_4 + SiO_2 \rightarrow Mg_2(Si_2O_6)$$

Chrysotile asbestos usually contains small amounts of brucite, $Mg(OH)_2$, which dehydroxylates at 400 °C. It loses strength when heated and it melts at 1521 °C.

2. *Uses*

The two unique properties of asbestos, its heat resistance and its fibrous nature, render it most suitable for the manufacture of heat-resisting textiles, e.g., garments and gloves used to protect workers exposed to high temperatures, and fire fighters. Textiles in which asbestos constitutes 75–99% can protect a person from temperatures of 200–480 °C. However, the consumption of asbestos in this area is only about 1%. The major use of asbestos is in the manufacture of asbestos cement. In this product, asbestos constitutes 12–15%, greatly increasing the resistance of cement as a construction material. Asbestos in admixture with plastics and asphalt features in the manufacture of a variety of tiles and sheets used as a covering material. Because of its heat resistance, asbestos mixed with metallic fibers and resins to the extent of 40% is used in the manufacture of what are called friction materials, e.g., brakes for motor cars and airplanes. Asbestos also features in the manufacture of certain industrial papers and gaskets.

3. *Structure*

There is a large variety of asbestos minerals, but the most important is chrysotile which represents more than 90% of the asbestos used; it is a hydrated magnesium silicate, $3MgO \cdot 2SiO_2 \cdot 2H_2O$. The given formula, however, represents only the ratio of the components and a more exact formula would be $Mg_3(Si_2O_5)(OH)_4$. To understand the properties of this material it is essential to know its structure. The basic building block of the silicates is the $(SiO_4)^{4-}$ tetrahedron. The variety in the structures of the silicates is due to the various possible combinations of these tetrahedra, both with each other through oxygen atoms and also with other ions. The negative charge on the tetrahedra is compensated by cations attached to the silicate structure by electrostatic forces, e.g., the single isolated tetrahedra, the double isolated tetrahedra, the rings, the chains, and the sheets. In the framework structure, the tetrahedra are connected together in three dimensions forming silica, a giant $(SiO_2)n$ structure.

Hydrated silicates are formed in nature by the weathering of anhydrous silicates. Serpentine is the rock that undergoes transformation to asbestos. The asbestos fibers that are visible to the naked eye are composed of hundreds of thousands of monofibers called fibrils in close-packed, parallel arrangement (Fig. 1). The monofibers have a tubular shape formed by the rolling up of a sheet composed of hydrated magnesium silicate. A typical fiber consists of 12–20 layers rolled up like a scroll (Fig. 2). The inner and outer diameters are in the order of 5–25 nm. The hydration represents about 10% of the mass and is in the form of hydroxyl groups linked to the magnesium and covering the external surface (Fig. 3). A sliver of chrysotile asbestos with a cross-section 0.1 mm square contains about 2×10^7 fibrils all in parallel orientation. It is possible, therefore, to strip from an asbestos fiber bundle very fine threads each of which still contains many thousands of fibrils.

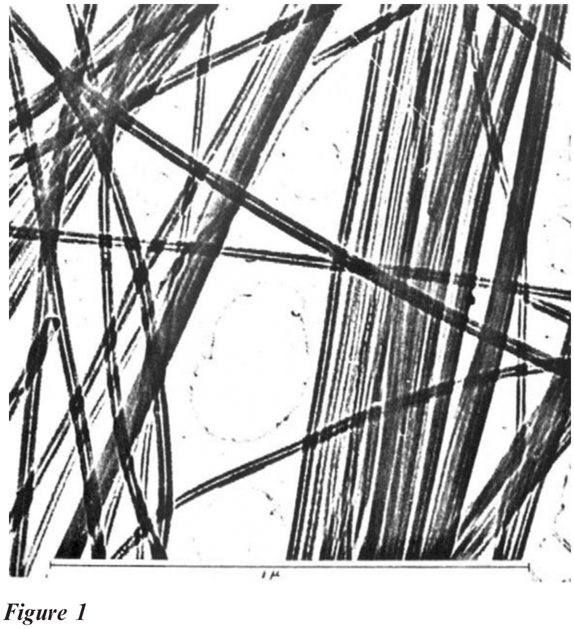


Figure 1
Asbestos fibers under the microscope.

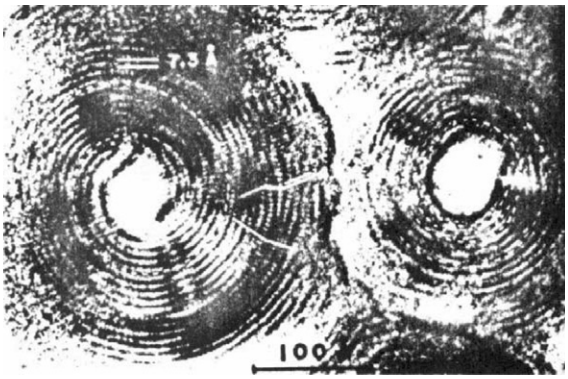


Figure 2
Cross-section of fibrils under a high-resolution electron microscope showing the scroll like structure.

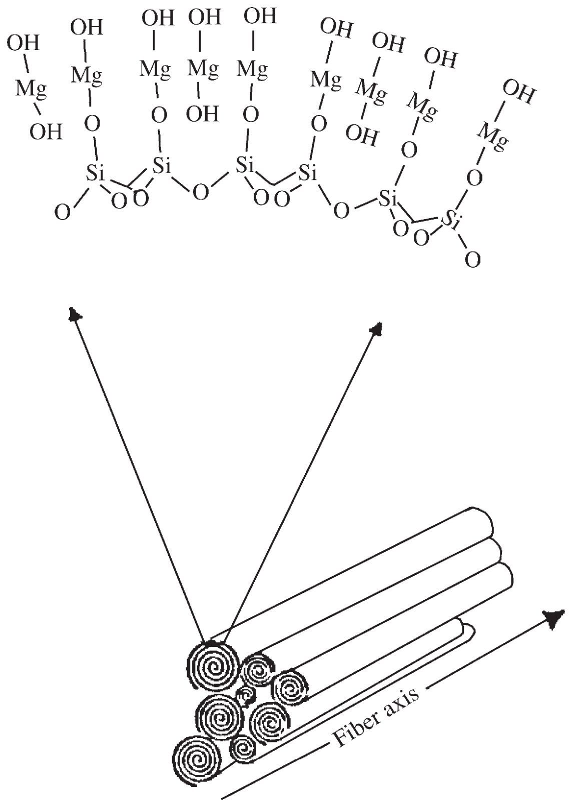


Figure 3
Schematic illustration of the hydrated magnesium silicate sheets rolled in the form of scrolls to form an asbestos fiber.

4. Problems and Suggested Solutions

There are two problems facing the asbestos industry in the early 2000s: its toxicity and the residues generated from its extraction. Both problems are under intensive investigation and potential solutions have been found.

4.1 Toxicity Issue

Asbestos is claimed by some to be a workplace problem, rather than an environmental issue, that can be adequately addressed with current dust control technology and appropriate workplace practices and by others to be a toxic and carcinogenic material that must be banned immediately (Selikoff and Lee 1978). It was demonstrated, for example, that heavy smokers in particular are more susceptible to have cancer of the lung; it seems that any fiber lodging in the lung because of its large surface area is capable of adsorbing carcinogenic substances from tobacco smoke. The accumulation of silicious dust in the lungs may result in causing silicose; in the case of asbestos this condition is known as asbestosis. But this may be cured by surgery. In the case of lung cancer due to asbestos, however, this is not possible because the disease spreads rapidly from the lungs to other organs.

As a result of this controversy, the consumption of asbestos in the USA, one of the world's largest consumers, has been decreasing gradually since the late 1990s. This has forced production decline and in some cases plant shutdown. The US industry is looking for substitutes for asbestos, but the long-term environmental effects of other minerals or synthetic substitutes that have been proposed are unknown. There are two widely accepted screening toxicity tests for asbestos:

(i) *Hemolytic*. In these tests, red blood cells are left in contact with a certain amount of fiber and after various periods of time the cells are centrifuged, and the hemoglobin content of the supernatant solution is measured using spectrophotometric methods. Usually, 72% of the cells undergo hemolysis when contacted with dry fiber, and 65% when contacted with wet fibers.

(ii) *Cytotoxicity*. In these tests, P $388D_1$ cells are incubated with a certain amount of fibers for various periods of time. The number of healthy and dead cells are then measured using Trypan Blue dye. Usually, 81% of the cells die when contacted with dry fiber and 50% when contacted with wet fibers.

While hemolysis is a short-term test in which the membrane stability is involved, cytotoxicity as outlined above gives a global picture of the cellular toxicity of the fibers. Hence, the two tests do not necessarily correlate, since the factors involved are not the same. Research effort is now directed towards understanding this toxicity problem and finding solutions. Two main routes have been so far proposed: reaction with phosphorus compounds, and reaction with chelating agents. These routes were chosen because it was observed that Belgian workers exposed to asbestos cement dust were exempt from developing cancer as compared to those exposed to asbestos alone or cement alone. The same phenomenon was confirmed by experiments carried out on rats at Battelle North West in the USA. It seems, therefore, that modifying the asbestos surface is a valid approach.

(*a*) *Reaction with phosphorus compounds*. Intensive research was conducted in Québec based on the treatment of the fiber with a dosed amount of gaseous phosphorus oxychloride, $POCl_3$, at about 200 °C to fix about 1% phosphorus on the fiber, thus rendering it nontoxic but maintaining all the desirable properties of asbestos. The asbestos so treated is known by the trade name *Chrysophosphate* (Khorami and Lemieux 1987). The reaction on the surface can be represented by:

$$3\ \mathrm{Mg(OH)_2}\ \text{in asbestos} + 2\ \mathrm{POCl_{3(g)}} \rightarrow \mathrm{Mg_3(PO_4)_2} + 6\ \mathrm{HCl_{(g)}}$$

Phosphorus oxychloride, however, is a reactive chemical that reacts with humidity in the air and requires special handling, and the Chrysophosphate releases phosphate and chloride ions when it comes into contact with water, e.g., in the manufacture of asbestos cement. Work at Laval University has shown that a similar product could be obtained by reacting the asbestos with an aqueous solution of ammonium phosphate, $(NH_4)_2 \cdot HPO_4$, at ambient conditions, followed by filtration and washing (Awadalla *et al.* 1990). The reaction on the surface can be represented by:

$$3\ \mathrm{Mg(OH)_2}\ \text{in asbestos} + 2\ \mathrm{HPO_4^{2-}}_{\ (aq)} \rightarrow \mathrm{Mg_3(PO_4)_2} + 4\ \mathrm{OH^-} + 2\ \mathrm{H_2O}$$

(*b*) *Reaction with chelating agents*. Certain chelating agents that react with magnesium ions in aqueous solution to form precipitates were found at Laval University to reduce the toxicity of asbestos (Habashi *et al.* 1991, Habashi 2001). The amount of chelating agent sorbed on the fiber is 2–5%. Chelating agents are organic compounds, usually with high molecular weight, that react with magnesium ions to form a ring structure through the lone pair of electrons on the oxygen or nitrogen atoms of the reagent. For example, the reaction of alizarin with magnesium ions forms an insoluble six-membered ring chelate; the covalent bond between the magnesium atom and oxygen of the quininoide structure is due to the lone pair of electrons belonging to the oxygen, while that between magnesium and the other oxygen of the hydroxyl group is due to equal sharing of electrons.

Asbestos fibers treated with chelating agents are vividly colored. The color has been shown to be stable on boiling with distilled water under a reflux condenser for 1 h, thus indicating that the organic reagents are chemically bound to the surface of asbestos and not physically adsorbed by the weak van der Waals forces. If they were physically adsorbed they should have been washed away by boiling. Under the microscope the colored asbestos samples were fully homogeneous and no precipitates were visible. This was also confirmed by examination with scanning electron microscopy. Appearance of new peaks in X-ray diffraction patterns and in infrared spectra, together with the absence of precipitates as revealed by scanning electron microscopy, supports the view that chelates are formed with asbestos, apparently with its $Mg(OH)_2$ component.

Chelate formation on asbestos fibers is similar to the formation of insoluble complexes of acid and basic dyes with metallic hydroxides precipitated on textile fibers by the process known as "mordant dyeing". In this process, the fabrics are soaked in a solution of aluminum or magnesium acetate, squeezed to drain excess solution, then heated in steam chambers to decompose the acetate and precipitate the metal hydroxide on the surface. The fabrics are then immersed in a dye solution whereby a colored chelate is precipitated on the surface known in the industry as a "lake" (Habashi 1992).

Among the dyes that were found to decrease the toxicity of asbestos were Thiazol Yellow G (a monoazo dye) and Trypan Blue (a diazo dye), but of course more dyes have to be tested to find the most effective and economic ones.

4.2 Stabilization or Utilization of the Tailings

As already mentioned, to recover 1 ton of asbestos, about 20 tons of waste rock must be removed. In the Province of Quebec alone, when 1.5 million tons of asbestos were produced annually during peak production, about 30 million tons of waste rock were rejected. More than 600 million tons of such waste have accumulated over the years. The problem with this waste is that it is now finely ground and, therefore, no longer stable as it is before mining, i.e., when wind blows, fibers and dust particles are scattered in the environment, which is undesirable. Regular attention is required to prevent excessive dust in the surroundings. Research has shown that the tailings can be stabilized by vegetation. Due to their high alkalinity, however, it would be necessary to add some sulfide tailings from a nonferrous metal mine or flower of sulfur to lower the pH to near normal values.

The tailings have practically the same chemical analysis as the asbestos fibers; their magnesium oxide content is about 40%. The following solutions have been proposed and tested and the results seem to be promising:

(i) Asbestos tailings represent an appreciable nickel reserve. Nickel is mainly in isomorphous substitution with magnesium; the remainder is in the form of a magnetic Ni_3–Fe alloy and the sulfide pentlandite. It is therefore possible to subject it to magnetic separation to obtain a concentrate rich in nickel and iron.

(ii) The tailings can be melted with carbon in an electric furnace to produce a molten silicate phase and an iron–nickel alloy containing 5% Ni. The silicate phase can be blown by air to form mineral wool for insulation.

(iii) Asbestos tailings can be leached to recover magnesium. The leach solution can be best purified by precipitating the impurities at pH 7 using MgO. The purified solution can then be subjected to crystallization. Hydrochloric acid is the preferred acid for leaching:

$$Mg_3Si_2O_5(OH)_4 + 6\ HCl \rightarrow 3\ MgCl_2 + 2\ SiO_2 + 5\ H_2O$$

because the crystals obtained can be dehydrated to get $MgCl_2$ for metal production, or decomposed to get MgO for refractories. When H_2SO_4 is used, $MgSO_4$ can be recovered and be utilized as a fertilizer. Another route suggested for producing metallic magnesium directly from the tailings is based on a process similar to the Magnetherm Process, i.e., the dehydroxylated tailings are mixed with ferrosilicon and heated in an electric furnace to vaporize magnesium which can be condensed and recovered.

(iv) When the tailings are calcined at 1300 °C in the presence of minor amounts of bentonite, a granular product suitable for use as foundry sand is obtained.

(v) The fact that serpentine loses its water when heated and the forsterite formed at high temperature reacts with silica present to form enstatite suggests that asbestos tailings may be used as a raw material for manufacturing refractory products.

Bibliography

Awadalla F T, Habashi F, Pagé A 1990 Reaction of chrysotile asbestos with phosphate ion in relation to toxicity. *J. Chem. Tech. Biotech.* **49**, 183–96

Habashi F 1988 Asbestos, progress and problems. *Arab Min. J.* **7**, 53–62; also in *Ind. Min. Eng. J.* **26**, 7–17

Habashi F 1992 Dyeing of asbestos fibers. *Textile Chemist Colorists* **24**, 23–5

Habashi F 2001 Chemisorption of organic dyes on chrysotile asbestos and the toxicity issue. *Eur. J. Min. Proc. Env. Protect.* **1**, 1–9

Habashi F, Awadalla F T, Pagé M 1991 Surface modification of chryotile asbestos with organic reagents. A preliminary *in vitro* toxicological study. *Bull. Can. Inst. Min. Met.* **84**, 67–79

Hodgson A A 1986 *Scientific Advances in Asbestos. 1967 to 1985*. Anjalena Publications, Crowthorne, Berkshire, UK
Khorami J, Lemieux A 1987 The phosphorylation of chrysotile asbestos fibers with phosphorus oxychloride ($POCl_3$): mechanism of reaction and chemical composition of the external coating. *Can. J. Chem.* **65**, 2268–76
Michaels L, Chissick S S (eds.) 1979 *Asbestos*. Wiley, New York, 2 Vols.
Selikoff I J, Lee D H K 1978 *Asbestos and Disease*. Academic Press, New York
Stewart P S B 1986 Asbestos cement manufacture from wet process chrysotile. In: *First International Conference on Asbestos Cement, Cannes, France*. The Asbestos Institute, Montréal, Canada
Virta R L 1999 *Asbestos*. In: *Minerals Yearbook*. US Geological Surveys, Washington, D C, Vol. 1, pp. 8.1–8.5

F. Habashi
Laval University, Quebec City, Quebec, Canada

Automotive Body Materials

The materials employed in high-volume mass produced passenger vehicles (cars and light trucks) are the result of a long evolutionary process begun at the beginning of the automotive industry and still continuing. Over time, the functionality and requirements for vehicles have increased. Today's vehicles must meet a wide variety of customer needs and societal and legislative imperatives, e.g., safety and crashworthiness, fuel economy, affordable cost, repairability, performance, styling, exhaust emissions, and recyclability. Materials influence or determine most of these characteristics and features and, therefore, selections of materials for automotive applications are part of design solutions, which can and do profoundly influence customer purchase decisions and the vehicle ownership experience. Many of the technical improvements in vehicles have been directly or indirectly enabled by developments in materials technologies. Fuel economy, emission, and safety improvements have been the key motivators for evolutionary material changes in the recent past. In the near future a combination of customer expectations, environmental stewardship, and tightened regulations will lead to significant changes in the technologies employed in automotive vehicles; materials will be at the forefront. For example, development programs around the world, such as PNGV in the USA and the 3-liter program in Europe, have goals of achieving quantum improvements in fuel economy and emissions are underway. Vehicle weight reductions, through the use of lower density materials such as aluminum, magnesium, and polymer composites to replace heavier materials are a key element in achieving these goals. The implementation of more efficient, alternate powertrain systems, such as hybrid-electrics or fuel cells, will also lead to significant changes in materials. Table 1 shows the trends in material selection for a number of major parts or subsystems of the vehicle.

Table 1
Vertical body panels—plastics and composites.

Component or subsystem	Current materials—widespread application	Aggressive material—implemented on some vehicles	Future potential—developmental
Body-in-white	Stamped steel, including high strength steel	Increased high strength steel usage; aluminum	Stamped aluminum; hybrid materials: aluminum, polymer
Body panels	Stamped steel	Stamped aluminum, SMC, thermoplastic moldings	Stamped aluminum, SMC, thermoplastic moldings
Engine block	Cast iron	Aluminum	Magnesium, elimination in favor of fuel cell
Transmission case	Aluminum		Magnesium
Suspension	Steel and cast iron	Cast aluminum	Magnesium, polymer composites
Brake system	Cast iron	Aluminum calipers	MMC aluminum rotors or drums
Wheels	Formed steel	Cast aluminum	Magnesium
Seat structure	Steel	Aluminum	Magnesium, polymer composites
Instrument panel beam	Steel assembly	Cast magnesium	Polymer composite

In order to be used in mass-produced consumer products such as automobiles and light trucks, materials must exhibit a number of general characteristics:

Plentiful and readily available supply of raw materials
Affordable cost
Fabricated at low cost at mass production volumes
Affordable tooling
Highly reliable parts and assemblies
Recyclability

These requirements tend to limit the choices and represent a significant (but appropriate) barrier to implementation of new materials. Changes in the materials used in production vehicles therefore tend to be evolutionary and based on the validated achievement of worthwhile benefits. Selection of materials for the various parts of the vehicle are based on the specific requirements, which in turn vary with vehicle size, type, regulations, and requirements for the geographic location of sale and use. The final selection of materials is, however, also influenced by the manufacturer's policy concerning such factors as pricing, purchasing, investment, employment, service, length of production series, and image. Because of this complexity of material selection, several different materials may be used in analogous parts for different vehicles or by different manufacturers.

Figure 1 shows the materials content of a medium-size passenger automobile, exhibiting what might be regarded as aggressive or state-of-the-art materials content for a current mass-produced vehicle. More typically, ferrous metals, sheet steel, forgings, and cast iron generally constitute up to about 70% of the vehicle by weight. Nonferrous metals together account for about 10% and plastics, both filled and unfilled, also account for about 10%. A passenger vehicle, automobile or light truck, has about 15000 parts. In the following sections the materials and manufacturing processes are described for the major subsystems of the vehicle. The selection has been made to illustrate the wide range of these materials.

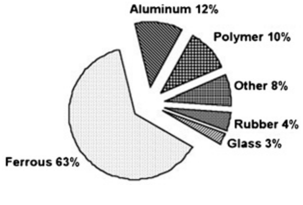


Figure 1
Typical materials content for a medium-sized automobile.

1. Steels

Automotive body components are currently manufactured predominately from ferrous materials, and this will likely continue for some time due to the advantages offered by ferrous materials such as low cost, formability, and recyclability. Most components are cold formed by stamping and then joined together by welding, usually spot-welding. The exterior body parts of the outer structure of a vehicle, such as panels, doors, and hoods, are coated for corrosion protection where necessary. The selection of a particular grade of steel is generally driven by either manufacturing requirements of the component such as formability, or by service requirements such as durability; the predominate requirement may limit the choice of grade. Formability is generally sacrificed for increased strength. Steel companies have remained competitive over the years by offering the ground vehicle industry a wide variety of products and improvements in the final quality through advances in processing. One of these advancements is continuous annealing. With the advent of continuous annealing, it became possible to develop steels with more uniform properties and lower alloy content that exhibit the same or better properties as box-annealed steels. Some of the main families of steels used in the ground vehicle industry are discussed below in general terms. Typical mechanical property values of the major automotive steels are summarized in Table 2.

Interstitial free steels contain carbon levels of about 0.004 and have the best formability characteristics of any grade of steel. The ability to routinely produce such steels is the result of advances in steel manufacturing. These ferritic structures have yield strengths of about 160 MPa and tensile strengths of about 300 MPa. Tensile testing may be used to assess formability through examination of the total elongation (%EL), the anisotropy parameter R, and the work hardening exponent n. For interstitial free steels these are 44%, 2.1, and 0.25, respectively.

Somewhat less formable are the mild steels that have been the standard for a number of years and are still a major body material. Mild steels encompass a wide variety of compositions and grades and are manufactured under less stringent composition and process conditions thus offering a cost advantage. Mild steels are all fully aluminum-killed, with the varying degrees of formability resulting from proprietary processing. Typically mild steels are ferritic in structure and have yield strengths of about 170 MPa and tensile strengths of about 310 MPa. Parameters for a deep drawing quality mild steel that reflect formability are 40% EL, R of 1.7, and n of 0.22.

Currently, mild steels are often being replaced by a relatively new grade of steel known as bake hardenable for exterior panels when dent resistance is a criterion. Bake-hardenable steels are made only on continuous annealing lines and as a family exhibit

Table 2
Mechanical properties of several steel grades.

Grade of steel	Yield strength (MPa)	Tensile strength (MPa)	Elongation (%)
Interstitial free			
IF	130	300	46
Bake hardenable			
BH210	227	345	37
BH280	280	380	31
Mild steel			
DDQSK	160	300	43
DQSK 1006	165	300	42
DQ 1006	190	300	41
CQ 1006	200	320	39
REPHOS			
DR	200	370	40
High strength low alloy			
HSLA 040XLK	300	390	29
HSLA 050XLF	380	455	27
HSLA 060XLF	440	530	25
HSLA 070XLF	510	570	23
HSLA 080XLF	580	640	21
Dual phase			
DP290-450	290	450	30
DP310-500	310	500	27
DP330-500	330	500	26
DP350-600	350	600	23
DP370-600	370	600	22

higher yield strength and similar formability compared to mild steel. Strengthening of bake-hardenable steels occurs through two mechanisms. The first is the obvious strengthening that occurs due to cold work, which occurs during forming. The second mechanism is by heat treatment of the steel, which occurs during the paint bake cycle which is applied to the entire automotive body. This strengthening occurs by carbon in solid solution preferentially locating to the generated dislocations and impeding further dislocation movement with subsequent deformation. Therefore, the higher yield strength is the result of both the heat treatment and cold work. Components that receive no deformation in the forming process, such as flat panels, would require some type of uniform strain in order for this grade of steel to be effective. A strain of about 2% and heating for 30 minutes at 175 °C will increase the yield between 55–82 MPa.

To obtain even higher strength, several different strengthening mechanisms are used. High-strength, low-alloy steels (HSLA) obtain their properties through careful thermomechanical treatment. This processing controls the grain size and precipitate distribution. The final properties result from the combination of grain size, solid solution strengthening, and a small contribution from precipitation hardening. The use of continuous-annealing lines has eliminated the variability of HSLA steels that was often found with box-annealed products.

Re-phosphorized steels are a grade of mild steel which offer increased strength while still maintaining reasonable formability. These steels are strengthened by phosphorous through solid–solution strengthening. Careful control over the sulfur and carbon content must be maintained in order to avoid spot welding problems.

Dual phase steels obtain their properties from a microstructure that contains a martensitic phase uniformly distributed throughout a ferrite structure. The strength of these steels can be varied significantly by varying the amount of martensite. The formable grades are characterized by a low yield-to-tensile strength ratio, high elongation, and work hardening. The amount of martensite is generally controlled by the amount of carbon, and the creation of the dual phase structure results from the presence of manganese in combination with appropriate choices of rolling practice and cooling.

1.1 Precoated Steel

The use of thinner, lighter weight sheet coupled with legislation that dictates corrosion standards make

Table 3
Typical mechanical properties of coated steels.

Grade	Yield strength (MPa)	Tensile strength (MPa)	Elongation (%)
Electro-galvanized			
Interstitial free			
VDIF	165	305	42
Mild steel			
DDQSK	165	300	43
DQDK 1006	180	305	42
Bake hardenable			
BK210	230	340	35
BK280	280	380	31
HSLA			
040XLF	303	390	29
050XLF	380	455	27
Hot-dipped			
DQ	250	345	35
CQ	260	360	33
HSLA			
HSLA 50	345	415	20
HSLA 60	415	480	16
HSLA 70	480	550	12

coated corrosion–resistant steel essential in many areas of a vehicle. For steels produced by continuous annealing, the reheating that occurs in hot-dip galvanizing may cause marked deterioration in properties. Therefore electro-galvanizing and zinc-rich paint systems are the only feasible means of corrosion protection. However, high temperature paint cure cycles can cause unacceptable stretcher strain problems or can change properties in some products, particularly those that depend on some type of mechanical abrasion such as shot-peening for their final properties. Table 3 lists some typical mechanical properties of coated steels.

In-line, hot-dip galvanizing produces the best potential corrosion protection, provided an adequate thickness of coating is achieved. Zinc acts as the sacrificial anode and is extremely effective as long as the protective layer remains intact. However, hot-dipped galvanizing can affect the final properties of sheet steel, particularly HSLA steels, by reducing both their strength and formability. The zinc layer can also interfere with spot welding by reacting with the electrode tip, causing the need for increased tip dressing.

Over the years, significant weight savings have been realized in automotive structures and components by vehicle downsizing and/or incorporating lighter weight materials such as aluminum, magnesium, and plastics. Weight savings are also being achieved with ferrous materials with judicious design and new technologies in the steel industry. Weight savings may be obtained through the use of higher strength steels (DP, HSLA), but the most effective method is to use them in conjunction with new technologies such as tailor-welded blanks and hydro forming.

1.2 Aluminum

Aluminum is increasingly used in body applications as a method of reducing vehicle weight. Table 4 lists some typical aluminum alloys used in automotive body components. Most current applications are 6000 (Al–Mg–Si) series alloys used for exterior panels, such as hoods, deck lids, and fenders. These are formed in the T4 temper and undergo precipitation hardening during the subsequent paint baking process thus increasing the yield strength to provide dent resistance in the finished panel. (Aluminum sheet producers have developed proprietary variations of the T4 temper, which enhance the aging behavior.) The use of aluminum for body structures is more developmental. Most manufacturers have developed prototypes, but only a few lower volume vehicles have reached commercial production. There are two main approaches to the design of aluminum vehicle bodies: (i) Monocoque construction (analogous to steel vehicle construction) consisting mostly of stamped sheet components joined by spot welding, riveting, and, in some examples, in combination with adhesive bonding. Bodies of this type are mainly composed of 5000 series (Al–Mg) alloys, which provide adequate strength and offer better formability and lower cost compared to other alloy types. Specific monocoque designs may also include the use of

Table 4
Examples of aluminum alloys for automobile body applications.

Alloy	Type	Yield strength (MPa)	UTS	Elongation (%)	Typical applications
Sheet (nonheat-treatable)					
5754	Al–3Mg	100	220	26	Body structure, inners for exterior panels
5182	Al–4.5Mg	130	275	24	Body structure, inners for exterior panels
Sheet (heat treatable)					
6111-T4	Al–Mg–Si–Cu	150 (230)[1]	280 (320)	26	Exterior panels
6022-T4	Al–Mg–Si	150 (215)	255 (290)	26	Exterior panels
6016-T4	Al–Mg–Si	125 (180)	235 (260)	26	Exterior panels
Extrusions					
6063-T6	Al–Mg–Si	215	240	12	Body structure components
6061-T6	Al–Mg–Si	275	310	12	Body structure components, bumpers, intrusion beams
7029-T5	Al–Zn–Mg	380	430	15	Bumpers, intrusion beams
Castings					
A356.0-T6	Al–7Si	280	207	10	Body structure Components

[1] Numbers in parentheses indicate typical strength levels, which might be realized in parts after 2% tensile prestrain (simulating forming strain) followed by 30 minutes of aging at 177 °C (simulated paint bake cycle).

castings, extrusions, or other product forms in selected components. Examples of monocoque aluminum vehicles include the Ford AIV and P2000, the GM EV1, and the Honda Insight. (ii) 'Space Frame' bodies in which the load carrying elements consist primarily of extrusions and castings with use of conventional stamped sheet more limited. Joining methods may include some combination of rivets, spot, MIG, and laser welding. The Audi A2 and A8 and the Ferrari 360, although different from one another, are vehicles based upon the space frame concept.

With respect to weight reduction, aluminum is an attractive body material, since reductions on the order of 50% compared with steel can be achieved. Although much progress has been made in the development of technologies for body applications it is still a higher cost alternative. Future growth in the number of applications will depend on achieving reductions in both material and manufacturing costs.

1.3 Plastics and Composites

Since the mid-1970s, the use of plastics and composites in automobile components has expanded from approximately 4% (by weight) of a typical car or truck in 1975 to more than 10% (by weight) in 2000. These materials can now be found across the entire gamut of vehicles produced by automotive industry including cars, light- and medium-duty trucks, sport utility vehicles, and mini-vans (see Table 5). Furthermore, they are used in a variety of applications including interior and exterior of vehicles, structural parts, and under-the-hood components. Of these applications, exterior components represent the most challenging use of plastics and composites because of the Class 'A' surface finish requirements and the fact that most automobile manufacturers attach these parts to the vehicle in the body shop and they must withstand the high temperatures associated with the paint shop's bake ovens. Nonetheless, plastics and composites are now commonly used for horizontal body panels (i.e., hood and deck lids), vertical body panels (i.e., front fenders, rear fenders, and door outer panels), and several emerging applications that require a combination of structural properties and surface finish.

(*a*) *Horizontal body panels*. Horizontal body panels represent the most demanding application of plastics

Table 5
Vertical body panels—plastics and composites.

Material	Process	Application	Vehicle
SMC	Compression molding	Front and/or rear fenders	Lincoln Continental Plymouth Prowler Ford F-150 Flareside Ford Ranger Splash Chevrolet C/K Pickup Truck
SMC	Compression molding	Door outer panels	Pontiac Firebird Chevrolet Camaro Chevrolet Corvette Saturn EV1
Polyurethane with mica or wollastonite filler	RRIM	Front and/or rear fenders	Pontiac Firebird Chevrolet Camaro Chevrolet Corvette Saturn EV1 Dodge Ram T-300 Chevrolet Silverado
PPE/PA blends	Injection molding	Front and/or rear fenders	VW Beetle Land Rover Freelander Mercedes A-Class Renault Clio II Renault Megane Scenic

and composites because of their location on the vehicle. This refers to both the performance requirements (i.e., part stiffness, fit and finish, thermal loads, and environmental degradation) and the appearance requirements since the customer constantly views these components whether they are driving their vehicle or simply walking past. To date, plastics have not been able to meet these requirements, however, through the inclusion of fibers and fillers in a neat plastic, thermoset composites provide all the properties necessary.

The composite material most commonly used for a horizontal body panel is sheet molding compound (SMC), which is processed by compression molded into its final form. To meet the structural requirements for hoods and deck lids, a two-piece design (i.e., inner and outer panels bonded together with an epoxy adhesive) is typically used. With this design, metal inserts, which are bonded and riveted to the inner panel, are used to attach the body panel to the rest of the vehicle. This approach leads to weight savings of 15–25% relative to steel. A short list of some of the vehicles that currently use SMC for horizontal body panels includes the Ford Mustang, the Lincoln Continental, and the Chevrolet Corvette.

Composite hoods and deck lids have also been fabricated by resin transfer molding (RTM); however, this process has only been used for low volume vehicles (i.e., <10 000 units per year) such as the Aston Martin DB7 and the Dodge Viper because of the long cycle times associated with this process and the amount of re-work required to achieve a Class ‘A’ surface finish.

(*b*) *Vertical body panels.* Vertical body panels offer a wide variety of material/processing combinations because their performance requirements (i.e., part stiffness, dimensional tolerance, etc.) are less demanding than those of horizontal body panels. Compression molded SMC, reinforced reaction injection molded (RRIM) polyurethanes and injection molded polyphenylene ether/polyamide (PPE/PA) blends have all been used to fabricate fenders for cars and/or light trucks. Regardless of the technology selected, fenders are manufactured as a complex shaped, single shell structure, which leverages the design/styling flexibility that plastics and composites offer. Table 5 contains a list of plastic and composite materials, processes, applications, and some of the vehicles on which they are used. Most of these parts are assembled onto the vehicle in the OEMs body shop, which means that the plastic or composite body panel must withstand the temperatures associated with the electro-coat, primer/surfacer, basecoat, and clearcoat bake ovens. In some cases, however, automobile manufacturers have made allowances for the plastic and/or composite material and either they attach the part after the electro-coat process or they slow down their assembly line and use lower bake oven temperatures. These strategies provide a mechanism for managing the thermal growth associated with RRIM polyurethanes and injection molded PPE/PA blends.

(*c*) *Emerging applications of plastics and composites.* Recent developments in the automotive industry such as the growth of derivative vehicles and the increased interest by OEMs in the automotive aftermarket provide new opportunities for plastics and composites. Foremost of these new opportunities is the use of composite materials for pickup truck box inner structures. Not only does the use of composites in this application eliminate the current steel box inner, but it also eliminates the plastic bedliner used to protect the steel from damage. Currently, two approaches exist for the fabrication of a composite pickup truck box inner: Ford is using compression molded SMC for the Explorer Sport Trac while GM is using structural reaction injection molding (SRIM) for the Chevrolet Silverado. Although a composite box inner structure does not require a Class 'A' surface finish, it must meet some appearance standards and all performance requirements including weathering and off-road durability.

Other emerging automotive applications for plastics and composites include removable hardtops for convertibles, rigid tonneau covers, and truck caps. These applications may actually be more complicated than composite pickup box inner structures because not only do they require similar levels of performance but they also all require a Class 'A' surface finish and their fit to the rest of the vehicle is extremely demanding. Because of these requirements, automotive OEMs have only used compression molding of SMC to fabricate these components for vehicles such as the Ford Explorer Sport Trac and the Ford Ranger.

(*d*) *Benefits and limitations of plastics and composites.* Perhaps the primary reason that plastics and/or composites are used to fabricate exterior components for low- to medium-volume vehicles (i.e., $<100\,000$ units per year) is the lower investment costs associated with this class of materials relative to steel stamping. For example, to fabricate a plastic or composite fender by RRIM, injection molding or compression molding requires only one mold cavity where as progressive stamping dyes are needed to stamp an equivalent part out of steel. However, at higher volumes, the reduced investment costs does not offset the higher variable costs associated with plastics and composites when compared to steel. Other benefits for plastics and composites relative to steel include reduced weight, improved corrosion resistance, improved dent/ding resistance, and greater design/styling flexibility.

In addition to the higher raw material costs of plastics and composites discussed above, there are some drawbacks to the use of plastics and composites for the fabrication of Class 'A' body panels. Surface defects such as pits and porosity can render a body panel unacceptable for the exterior of a vehicle. Other defects associated with Class 'A' plastic and composite panels include delamination, bond line read thru, sealer redeposition, surface waviness, and handling related problems such as cracking.

A. M. Sherman, A. R. Krause, P. A Friedman,
D. A. Steenkamer and D. Q. Houston
Ford Motor Company, Dearborn, Michigan, USA

Automotive Chassis/Suspension Materials

This section will examine the large number of elements connected to the vehicle frame that steer and stop the vehicle and support/suspend/isolate the body (and the passengers) from road vibrations. These elements are collectively termed chassis components and include the three vehicle subsystems: suspension, steering, and brakes. These three systems all interact to affect the ride and handling of the vehicle.

The suspension system is composed of five main subsystems: front suspension, rear suspension, shock absorbers, load leveling, wheels and tires. The steering system is composed mainly of vehicle interior elements (steering column, steering wheel, power steering subsystem); but there are a few chassis-mounted components (steering linkages) in this group. The brake system includes both a front and rear drum or disk subsystem, and other power/electronic elements that are not discussed here since they are not part of the overall vehicle handling (braking, steering, vibration) experience.

The front and rear suspension elements are only subtly different in their responsibility for supporting the front and rear of the vehicle. The main difference is that the front has a cross member, also called the subframe or cradle, while the rear has a unique rear cross member. Components that are more-or-less similar in the front and rear include the right/left upper and lower control arms, right/left knuckles/spindles, springs, shock absorbers, wheels and tires, and sway/stabilizer bars; in some classifications, the rear differential and nondriven axles and hubs are also included as chassis members. Within each element are complicated assemblies of elastomeric bushings that help isolate the vehicle body from undesired vibration and shocks without affecting the unique steering, ride, and handling characteristics that define how the manufacturer differentiates each vehicle. Wheels and tires perform both a suspension as well as steering and handling function. Front and rear brake components influence vibration and handling through two- and four- wheel disk brake rotors/drums, calipers, and friction materials. Master cylinder, brake booster assembly, pedals and brackets control the braking action but are not discussed here as part of the suspension–chassis package.

1. Design Considerations

Suspension characteristics are defined by the assembly of suspension components, and are designed by marketing and sales people, who try to translate customer wants into unique and saleable vehicle dynamics. Engineers try to translate the customer wants into definable suspension kinematics and compliance characteristics. In a recent article by GM engineers on the 1997 Corvette, Fuja *et al.* (1997) indicated that there were 60 suspension attributes or design factors that had to be monitored analytically in the development of the 1997 Corvette's overall vehicle dynamics, handling, and performance. As with any complex system, modifications that improve one particular suspension attribute may cause degradation in another; this will often require additional redesigns of all the elements in the system to ensure that the desired vehicle response has been achieved. In addition to vehicle dynamic constraints, there are often packaging and cost issues which impact on the final suspension design selected.

All chassis and suspension systems are now designed using extensive computer modeling. For example, cornering forces, and transient roll behavior can be predicted to be affected by shock absorber structure, spring effects, and other suspension elements. If there are 60 elements in a suspension system, there are probably a similar number of attributes in the braking system and wheel and tire system that can affect vehicle dynamics.

2. Introduction to Materials Selection Criteria

The materials selected for chassis and suspension components differ from company to company, since each is trying to develop a unique product for its perceived market. Any change in design will interact with the materials employed, and the manufacturing process used to fabricate each component will be affected by the material choices. This is true of the individual metallic components as well as the rubbers and elastomeric materials used; all changes will cause a noticeable change in the vehicle's dynamics.

It is almost impossible to describe and classify the large number of materials employed in the various automotive chassis and suspension components in the ~50 million vehicles that are produced annually around the world. We have seen that each unique design will employ a unique suite of materials and fabrication process to achieve unique performance, as required for the vehicle's market niche. An important selection criterion is based on the fact that each component must satisfy its expected service conditions (straight driving on flat surfaces and over cobblestones, cornering, braking, accelerating over rough roads, sliding against curbs/potholes, etc.) and must satisfy some expected durability condition (primarily a fatigue lifetime of 10^7-10^8 cycles versus a defined stress loading and stiffness criteria) with the required reliability and safety. Another issue that constrains material choices is whether the component will be installed on an automobile (usually of 1700 kg mass or less), or a truck (where the mass can exceed 3000 kg). In the latter case, the duty cycle is more aggressive, the loads are higher, and there is more concern about very long-term durability (>400 000 km); automobiles are usually expected to survive about 200 000 km.

Each material can be fabricated into a component in many ways, depending on the alloy selected, the shapes required, the section thickness(es), and the packaging (space available). As an example, consider a cast lightweight (aluminum) control arm. There are only a couple of alloy choices, viz. Al 356 or Al 357, out of which control arms are cast. And while there are seven ways to produce a cast control arm (gravity, low pressure, tilt, VRC/PRC (Alcoa), squeeze, semi-solid, and high pressure), there is a wide range of heat treatments: annealing-soaking temperatures and times, quench medium/quench temperature, and aging (and preaging) times/temperatures. These combinations are selected by individual foundry preferences to achieve the required tensile strength, fatigue strength, and ductility necessary for the component's durability.

Material selection is not a standardized process. Each company has a set of preferred suppliers who are authorized to fabricate desired components based on the complex set of criteria that define best-in-class function. The choice of material is not as important as is the demonstration of functional excellence, component durability, and of critical importance, the component price. In Table 1 the reader will see a set of preferred materials' choices for the major suspension and chassis components.

The criteria for future materials selection are based on five current pressures that are promoting technological change in the automotive industry: worldwide legislation to reduce automotive emissions; continuing increases in the price of petroleum-based fuels; improvements in vehicle safety; competitive requirements to reduce noise, vibration, and harshness (NVH) and improve vehicle handling; ability to reduce vehicle cost while adding value. These criteria essentially predicate ultra lightweight materials: new steels (high-strength low-alloy (HSLA) steels), lower-cost cast irons (compacted graphite (CG) irons), and aluminum and magnesium components. New computer-control-led processing techniques (e.g., hydroforming, laser welding, computer-controlled casting, etc.) are being developed to fabricate these new materials into low cost, durable, high-strength components and assemblies.

Many components, such as wheels, sway bars, front and rear subframes, shocks, and springs have historically been produced from steel fabrications.

Table 1
Materials used in suspension and chassis for trucks (T) and automobiles (A).

Component	Pre 1990		1990–2000		2005+	
	T	A	T	A	T	A
Subframes	fabricated steel		steel fabs	aluminum fabs, cast Al	steel, HSLA steel, cast Al	
Springs	steel		Steel		steel, HSLA steel, reinf plas	
Control arms	fab steel	SGI	fab steel	cast Al (50%)	steel, HSLA steel, cast Al	
Knuckles	SGI		SGI	CGI, SGI, cast Al	Cast Al	
Shocks	fabricated steel		fabricated steel		fab steel	Ti, steel
Wheels	steel	cast Al	cast Al, steel	cast Al, steel	cast Al, steel	cast Al, Mg
Brake rotors	GCI		GCI		GCI	GCI MMC Al
Brake drums	GCI		GCI	GCI, MMC Al	GCI	GCI, MMC Al
Calipers	SGI		SGI	GCI, Al, MMC Al	SGI	Al, MMC Al

SGI, spheroidal graphite or ductile cast iron; CGI , compacted graphite cast iron; GCI, gray cast iron; reinforced plastic, composite of polymer with fiber glass; MMC Al, composite of aluminum with SiC particles; SSM Al, semisolid molded or cast aluminum; fab steel, fabricated, usually by welding steel tubing; HSLA steel, high-strength low-alloy steel.

But there has been a gradual increase in the use of aluminum for lightweight replacements. Aluminum subframes are being fabricated from extrusions, stampings, and castings, with a beginning use of hydroforming technology. Cast aluminum wheels now make up over 50% of road wheels in cars and light trucks; forged wheels are used at the higher end (such as the 2002 GM Corvette). Cast magnesium wheels have been used on racing vehicles, and as magnesium ingot prices drop and R&D continues to grow, we can expect a continuing presence of this ultra lightweight metal in automotive components. Suspension control arms currently are being produced from steel, ductile iron castings, and, recently, aluminum castings. Many steering knuckles have been produced since 1970 only as ductile iron castings (prior to being a malleable iron casting); recently there has been an increase in the number of vehicle lines using cast aluminum knuckles, produced by the semisolid molding process as well as by pouring into conventional permanent molds. One or two vehicle lines have used compacted graphite cast iron due to its lighter weight and lower cost. Springs and shocks play an important role in vehicle stability. The new steels and associated processing required to develop lightweight springs requires higher hardness steels ($>52\,\mathrm{Rc}$) and the associated higher strength to $>1200\,\mathrm{MPa}$ which can satisfy the high corrosion fatigue criterion that dominates this application (Yoneguchi *et al.* 2000).

Disk and drum brakes are made by casting a gray iron shape and then machining to a very high tolerance of flatness and parallelism ($12\,\mu\mathrm{M}$). But there has been considerable R&D to improve the iron's strength, as well as to increase its thermal conductivity, to increase overall heat transfer and thereby reduce brake temperatures; this will reduce distortion, and improve the vehicles' NVH characteristics. One vehicle line in both Germany and North America has seen the introduction of cast aluminum MMC for additional weight-reduction opportunities.

The growing demand for lighter-weight solutions is anticipated to consume a larger and larger fraction of a vehicle's total materials content. In addition to aluminum forgings, extrusions, and castings, high-strength steel fabrications, magnesium castings, and polymer composite assemblies are beginning to appear in production components. The critical issue is the ability to demonstrate lightweight designs that have good function, while satisfying long-term (200 000 km) durability and have a lower cost. New materials, component designs, and assembly processes are being developed continuously by the automotive industry. To understand the changing strategies of chassis and suspension components, the interested reader can read the references below, to see the work that has been recently performed and published in the US Society of Automotive Engineers (SAE) transactions.

Bibliography

Fuja S P, Schmidt H A, Ryan J P 1997 Paper #970097. Society of Automotive Engineers, Transactions. SAE, New York

Society of Automotive Engineers 1997 Key design developments of the 1997 Chevrolet Corvette. SAE-SP-1282. SAE, New York

Society of Automotive Engineers 1997 Steering and suspensions technology. SAE-SP-1223. SAE, New York

Society of Automotive Engineers 1998 Tire, wheel, steering and suspension technology. SAE-SP-1338. SAE, New York

Society of Automotive Engineers 2000 Brake technology. SAE-SP-1537. SAE, New York

Society of Automotive Engineers 2000 Computer applications for crash, optimization, and simulation research. SAE SP-1496. SAE, New York

Society of Automotive Engineers 2000 Steering and suspension technology symposium 2000. SAE-SP-1519, SAE. New York

Yoneguchi A, Schaad J, Kurebayashi Y, Ito Y 2000 Paper #2000–01-0098. Society of Automotive Engineers, Transactions. SAE, New York

G. Cole
Ford Motor Company, Dearborn, Michigan, USA

B

Bearing Materials: Plain Bearings

This article is concerned with the materials which form the working surface of plain bearings. They are most commonly non-ferrous alloys, but may be polymers or polymer-based composites or, rarely, ceramics. Rolling-element bearings and the properties of the steels used in their construction are outside the scope of the present article.

Most plain bearings are oil- or grease-lubricated, and among the major factors determining the choice of bearing material for a particular application are the type, amount, and quality of the lubricant. Some polymer-based bearing materials and composites incorporating solid lubricants have been formulated so as to operate without lubricant.

1. Crankshaft Bearing Materials

The major impetus to development of bearing materials over the second half of the twentieth century has been the internal combustion engine. Depending on the number of cylinders, the crankshaft is carried by between three and thirteen pairs of half bearings (the main bearings) and there are three to twelve pairs of rod bearings housed in the connecting rods and operating against the crankpins. World production of crankshaft half bearings for the various forms of internal combustion engine is something like one billion each year.

The increasingly severe operating conditions which new designs of engine have imposed on the crankshaft bearings have led to the virtual disappearance of the once familiar tin- and lead-based Babbitt alloys, and to their replacement by a range of stronger alloys based on aluminum or copper.

1.1 Aspects of Crankshaft Bearing Performance

(*a*) *Fatigue strength.* The crankshaft bearings of the internal combustion engine are hydrodynamically lubricated; that is, oil is supplied to them under pressure, and an oil film separating the bearing and the crankshaft is generated by the rotation of the crankshaft. Very heavy loads are applied to the bearing surface through the oil film, the two components of the load being the firing load, applied through the piston and connecting rod as each cylinder of the engine fires, and the inertia load, resulting from rotation of the piston and connecting rod relative to the crankshaft.

It is the prime requirement of the bearing that it should be able to carry these loads without suffering damage. The common crankshaft bearing has a steel backing with a thin lining of bearing alloy, and to some extent it is this construction of the bearing which provides its load-carrying capability. However, the strength of the lining alloy is also important, and in choosing an alloy for a particular application the first step is to compute the maximum value of the dynamic pressure generated in the oil film, and to ensure that the alloy has the strength to survive. The choice is made by experience, since there is no simple relationship between any mechanical property of the lining measured in the laboratory and the ability of the lining to provide the necessary fatigue life in engine service. If an alloy of insufficient strength is chosen, fatigue cracks will eventually appear in the lining, pieces of the lining will be lost and the hydrodynamic oil film will be disrupted.

(*b*) *Surface properties.* If there were no other considerations, the strongest alloy available would be chosen and would perform satisfactorily. However, other factors make it undesirable to choose an alloy with a higher fatigue strength than necessary.

Firstly, the bearing and crankshaft are not always separated by a continuous oil film. Imperfect crankshaft geometry, deflections under load, and stopping and restarting all lead to some degree of contact between the crankshaft and bearing. Rubbing takes place under high load and at high speed, and high surface temperatures are generated, which in extreme instances can lead to seizure—friction becomes so great that the crankshaft can no longer rotate in the bearing. Crankshaft bearing alloys are formulated for good compatibility, that is, a minimum tendency to weld to the crankshaft in areas of rubbing. Most crankshaft bearing alloys contain a soft, low-melting-point phase which smears over the bearing surface wherever high temperatures are generated, and thus seizure is prevented.

Dirt entering the bearing clearance with the oil flow is another source of bearing damage. If the dirt cannot be fully embedded into the bearing surface, it may score the crankshaft (depending on the relative hardnesses of the dirt and the crankshaft surface), and the roughened crankshaft may in turn cause wear of the bearings. Embeddability, which is inversely related to the hardness and strength of the bearing lining, is thus an important property.

When dirt is embedded, particularly a relatively large particle of in-built dirt such as may be left in crankshaft drillings, bearing material is displaced around the particle and appears as a ring of alloy protruding from the bearing surface. This ring is rubbed by the crankshaft, but again the presence of a soft, low-melting-point phase prevents the local welding or "pickup" of bearing alloy onto the crankshaft and avoids seizure.

---

* Cross references marked by an asterisk are included in this volume.

# Cross references marked by a hash can be found by consulting the Encyclopedia of Materials: Science and Technology.

Since the introduction of nodular cast iron crankshafts in passenger car engines, the crankshaft bearings have been called upon to perform another function, that of "conditioning" the crankshaft. Even after grinding and lapping, the crankshaft surface may contain around the graphite nodules ferrite "caps," which stand proud of the surface and can dig into the bearing during the thin oil film part of the operating cycle. Bearing wear and even seizure can result. In unplated aluminum bearings the incorporation of silicon, present as fine, hard particles uniformly distributed throughout the bearing lining, provides a polishing action, removing the ferrite caps and creating a non-aggressive crankshaft surface.

(*c*) *Corrosion and erosion*. Corrosion resistance and cavitation erosion resistance are another two important properties of crankshaft bearing alloys. Corrosion resistance is important because the trend to higher engine temperatures and longer oil-change intervals results in eventual degradation of the engine oil, and the appearance of an acidic component which can attack the crankshaft bearing. The lead phase of copper–lead bearings is prone to this form of corrosion.

Cavitation of the oil film results from the rapid movement of the crankshaft within the bearing during firing, or from other causes, leading to a disruption or a disturbance in the oil flow at a particular location in the bearing and a particular point in the engine cycle. Subsequent collapse of the vapor bubbles on the bearing surface leads to classic cavitation erosion. Erosion patterns generally show a recognizable symmetry, and in cross section the damage shows a distinctive undercut appearance. Cavitation erosion resistance of bearing alloys is broadly related to alloy hardness, but microstructure is also important. Tin-based whitemetal has a much higher cavitation erosion resistance than lead-based whitemetal of approximately the same hardness, probably due to the copper–tin needles in the microstructure.

(*d*) *Choice of alloy*. The foregoing considerations show that the choice of an engine bearing alloy is a compromise. Conformability—the ability to conform to a misaligned counterface—is inversely related to hardness, as is dirt embeddability, and to some extent also compatibility, because of the desirability of a soft, low-melting-point phase. On the other hand, fatigue strength and cavitation erosion resistance tend to be higher the harder the alloy. For many engines, a satisfactory compromise is found in a two-layer bearing construction, that is, a lining 0.2–0.5 mm thick of one of the alloys described in the following sections, bonded to a low-carbon steel backing. The alloy is chosen which offers the balance of bearing properties most suited to the requirements of the particular engine under consideration.

Recently it has been found that the conformability of a relatively hard bearing alloy may be improved by the creation in the surface during final boring of circumferential microgrooves, which provide a degree of "give" in areas of crankshaft contact.

For the more severe engine applications, a three-layer construction is adopted. On top of the bearing lining a thin layer of a soft, lead-based alloy is electrodeposited as the last operation in bearing manufacture. The overlay is necessarily thin because of the high engine loads applied to it; it is commonly 0.015–0.025 mm.

1.2 Lead- and Tin-based Whitemetals

The lead- and tin-based whitemetals, or Babbitts, are alloys of lead, antimony, and tin, or tin, antimony, and copper. The most popular lead-based alloys are PbSbl0Sn6 and PbSbl5SnlAsl; the most popular tin-based alloy is SnSb8Cu3. (Formulae of this type are used in this article to indicate compositions as recommended by the International Standards Organization: SnSb8Cu3 = Sn–8%Sb–3%Cu.) The former consists of cuboids of antimony–tin compound SbSn in a pseudoeutectic of lead–antimony–tin. The latter consists of needles of copper–tin compound Cu_6Sn_5 in a tin–antimony solid solution. Tellurium has a compound refining effect on the tin-based alloys; arsenic refines the antimony precipitate in PbSb15Sn1As1.

As a class, the whitemetals have excellent surface properties: compatibility, conformability, and dirt embeddability. However, their fatigue strength is inadequate for the majority of internal combustion engines, and these alloys survive only as crankshaft bearing linings in slow-speed marine diesel engines. The tin-based alloy mentioned above, often with a 1% cadmium addition for increased strength, is the most popular alloy for marine diesel applications.

1.3 The Copper–Lead Series

When the whitemetals began to show inadequate fatigue strength, attention was turned to copper-based alloys and a range of copper–lead and lead–bronze alloys was developed. At the upper end of the lead range, that is 40–50% lead, the surface properties were found to be sufficiently good for the bearings to be operated without an overlay, and for many years the so-called "intermediate" alloys were found in the majority of passenger car engines manufactured in the USA. The term "intermediate" indicated that the alloys lay between whitemetal and those overlay-plated copper–lead alloys with a lower lead content adopted for heavy-duty engines.

The demise of the intermediate copper–lead bearing was the result of corrosion of the lead phase as engine temperatures increased, and as adoption of lead-free fuel resulted in greater corrosivity of the engine oil.

The copper–lead alloys now in use do not exceed 30% lead and contain up to 4% tin. All are overlay-plated with lead–tin, lead–tin–copper, lead–indium, or lead–tin–indium, and it is one of the functions of the overlay to protect the underlying alloy against corrosion of the lead phase.

The microstructure of the copper–lead bearing alloys shows distinct copper and lead phases, there being virtually no lead in the copper phase or copper in the lead phase. The tin, if incorporated in the alloy, is associated with the copper phase, where it has a strengthening effect. Thus alloys such as CuPb14Sn3 or CuPb23Sn3 are capable of carrying higher bearing loads than CuPb23Sn1 or CuPb30. The higher-tin alloys are found in the most arduous applications, such as the connecting rod bearings of the high-speed diesel. Their higher hardness is compatible with the hardened crankshafts found in these engines. The CuPb14Sn3 alloy produced by a sinter process, modified to refine the lead phase, is the strongest of the overlay plated copper–lead alloys.

Still higher strength has been achieved in Germany by the application of physical vapor deposition to bearing production. An aluminum–tin overlay is applied to the surface of a copper–lead bearing by cleaning the surface in vacuum by ion bombardment, followed by the deposition by sputtering of 16 microns AlSn20. "Spütterlager" have been adopted by European engine builders for the most heavily loaded connecting rod bearings in truck and passenger car diesel engines, where the high cost of the bearings can be justified.

1.4 The Aluminum Series

Crankshaft bearing production worldwide is roughly equally divided between copper–lead and aluminum alloys, although the aluminum series covers a wider range both of alloys and engine type.

The aluminum crankshaft bearing alloys are corrosion resistant and do not therefore rely on the overlay to protect them from corrosion. Thus in passenger car engines, unplated aluminum-lined bearings are found in Europe, the USA, and Japan.

In Europe, reticular tin–aluminum, AlSn20Cu1, developed in the UK in the late 1950s, proved very successful and became well established. The term "reticular" refers to the network of tin islands, interconnected along trigonal grain boundaries, distributed throughout the aluminum–1% copper matrix. In Japan, derivatives of reticular tin–aluminum containing additions of antimony or of silicon, lead, and chromium have been developed.

In the USA, a wear problem was found with this alloy, associated with the surface finish obtained on nodular cast iron crankshafts in the early stages of their development, and aluminum–lead took the place of aluminum–tin. The lead content of aluminum–lead is 4–8%, and there is a small tin addition of 0.5–1.5% associated with the lead phase.

The alloy also incorporates a 4% silicon addition, providing the crankshaft polishing action referred to above, and avoiding bearing wear. Minor additions of copper and magnesium or manganese are made to the alloy for increased fatigue strength.

Early versions of the aluminum–lead–silicon alloy were made either by a low quench rate, continuous-casting process or by a powder metallurgy route. Both showed a poor microstructure associated with the metallurgy of the aluminum–lead system, and in consequence a less-than-ideal fatigue strength. More recently, a high quench rate continuous-casting process for the alloy has been developed, which results in a much finer lead phase. Heat treatments have been developed to grow the silicon phase to the optimum size for its crankshaft polishing action.

In Japan, aluminum–tin–silicon alloys were developed with the tin content reduced from 20 to 12%, and a 2.5% silicon content introduced together with 1.5–2.0% lead. Copper 0.7–1.0% was incorporated, together with other minor alloying additions for improved fatigue strength. Similar alloys have been introduced in Europe and the USA. In the USA, the high quench rate casting process has been used to produce an AlSn8Si2.5Pb2 alloy with copper and chromium strengthening additions. In the UK, AlSn10–12Si4Cu1 alloys have been developed, the alloys being given a solution heat treatment so that additional strengthening is achieved during age-hardening in the engine.

Aluminum–tin–silicon is now the most popular alloy worldwide for NCI crankshaft passenger car engines. It combines good fatigue strength, crankshaft polishing action, and seizure resistance, with the added advantage of being almost or totally lead-free.

Unplated aluminum–silicon has also been used successfully in medium-speed marine engines. However, for the more highly loaded crankshaft bearings in high-speed diesel engines, stronger overlay-plated aluminum alloys are required. The early low-tin alloys AlSn6Ni1Cu1 and AlSn6Si1.5Ni0.5Cu1 are still in use, but cadmium-containing alloys AlSi4Cd1 and AlCd3Mn1.5Cu1Ni1 have been dropped on environmental grounds. A cadmium-free version of the latter alloy is operating successfully in the USA, together with a slightly stronger alloy AlSi5Sn2Cu1Mn1Ni1.

Comparable strength aluminum–zinc–silicon alloys have been developed in Japan. All are overlay plated with lead–tin or lead–tin–copper on a thin copper interlayer. The copper interlayer is preferred to nickel by one major engine builder on anti-seizure grounds.

At the other end of the spectrum a soft aluminum–tin, AlSn40, has been developed in the UK specifically for the marine diesel engine. Conformability and seizure resistance are of paramount importance in these large engines, where the crankshaft and cross-head bearings are between 400 and 900 mm in diameter. Bearing seizure can result in a crankcase explosion and is to be avoided at all costs. This alloy has a hardness comparable with that of tin-based white metal, without the loss of fatigue strength at engine temperature associated with the latter.

1.5 Lead-based Overlays

The lead-based overlays on the crankshaft bearings of heavy-duty engines have to survive under very arduous conditions. The overlay has to be soft to provide good conformability and dirt embeddability. At the same time, it has to withstand oil film pressures which may be as high as 350 MPa.

The relationship between overlay thickness and fatigue strength is well established. For some engines, adequate overlay strength is obtained from an overlay 0.025 mm thick, although for engines with particularly high rod bearing loads an overlay thickness of 0.015 mm is preferred.

Unfortunately, none of the overlays in current use is completely stable at engine temperatures. Changes in the chemical composition of the overlay take place as a result of diffusion phenomena, and there is accompanying deterioration in both strength and corrosion resistance.

The loss of tin by diffusion from a lead–tin (PbSn10) overlay into the underlying copper–lead may be delayed by the addition of 1–5% copper to the overlay composition, and by the incorporation of a nickel barrier between the overlay and the lining. Although tin and copper are still lost from the overlay to form a ternary nickel–tin–copper compound on top of the barrier, sufficient tin is retained in the overlay to provide adequate bearing performance.

The higher-copper overlays are harder and have somewhat enhanced wear resistance: recently a change from a fluoroborate electrolyte to methane sulfonic acid has enabled both tin content and copper content to be increased. The resultant PbSn14Cu8 overlay has good wear resistance without loss of fatigue strength, and is more stable at engine temperature than other lead-based overlays, presumably as the result of copper–tin compound formation within the overlay.

Lead–indium overlays can be produced by sequentially plating the lead and indium in separate baths, and diffusing the indium into the lead by heat treatment. Under certain circumstances they offer higher fatigue strength than lead–tin, but suffer from similar diffusion problems. At engine temperature the indium diffuses into the lead phase of the copper–lead, resulting in a loss of overlay strength and of wear and corrosion resistance.

In Japan, improved wear and fatigue resistance has been obtained by the incorporation of both tin and indium in the overlay, in combination with a nickel interlayer to limit diffusion. Lead–tin–indium and lead–tin–indium–copper overlays have been adopted by Japanese engine builders.

The foregoing refers to overlay-plated copper–lead. For electrochemical reasons, lead-based overlays cannot be applied directly to aluminum. It is therefore standard practice to deposit an interlayer between the overlay and the alloy, similar to the nickel barrier used on copper–lead. The interlayer may be nickel or copper. At engine temperatures, diffusion of the tin into the interlayer takes place and there is the same danger of loss of overlay strength and corrosion resistance as exists with copper–lead bearings.

In view of the environmental problems associated with lead, there are ongoing developments among engine bearing suppliers aimed at a lead-free overlay. Tin-based compositions are being explored, as are high-temperature polymers incorporating a solid lubricant. As yet, none has reached the stage of large volume production.

2. Bearing Materials for General Engineering

In the varied category of general engineering applications, the degree of lubrication can be anywhere between hydrodynamic and boundary, the former implying complete separation of the surfaces in relative motion by an oil film, the latter implying continuous rubbing contact. Wick lubrication, splash, or oil mist lubrication and grease lubrication are intermediate forms.

The bearing conditions in this category are generally less severe than those imposed on crankshaft bearings; in the hydrodynamically lubricated applications, oil film pressures are lower, as are oil temperatures. Fatigue strength and corrosion resistance of the bearing materials thus become of less importance. On the other hand, in the nonhydrodynamic applications, wear resistance assumes a much greater importance than it does in crankshaft applications.

The majority of alloys found in this category are still copper-based, both steel-backed and solid; but steel-backed aluminum alloys are found in an increasing number of automotive applications. Lead- and tin-based whitemetals are common. Many polymers behave well in lubricated applications, and when the lubrication is sparse, polymer-based bearing materials often outperform metallic bearing alloys.

Oil-impregnated bushings in which the bushing wall is a lubricant reservoir, and which require no external application of lubricant, are included in this section. However, composites incorporating solid lubricants are considered in Sect. 3, even though they are sometimes used in conjunction with conventional oil or grease lubrication.

2.1 Copper Alloys

The lining alloy specifications for many steel-backed bushings and thrust washers are similar to those of the half bearings dealt with in Sect. 1.3. CuPb23Sn3 is a general-purpose alloy found as the bushing lining in a wide variety of applications, such as automotive automatic transmissions, refrigeration compressors, and hydraulic gear pumps.

Applications requiring a higher-strength lining use a lower-lead, higher-tin alloy, CuPb10Sn10. This alloy performs well in the heavily loaded piston pin bushing of the internal combustion engine. The high loads and oscillating motion of this application make the creation of a satisfactory oil film difficult, and a surface finish of 0.1 μm R_a is required on the hardened piston pin if an acceptable lifetime is to be obtained. Recent development trends in engine design are further increasing pin bushing loads and temperatures, resulting in corrosion by oil additives. A high-strength aluminum–bronze lining alloy CuAl8 meeting these conditions is now well established. Other lead-free, high-strength alloys such as copper–tin–nickel are under evaluation. Solid (non-steel-backed) brass bushings of composition CuZnSi1 or CuZn23 Al6Mn4Fe3 have been successfully introduced in some engines.

Solid bronze bushings are used in many applications where quantities do not justify the tooling costs associated with the forming of steel-backed bushings and a suitable stock size is not available. The bronzes are usually continuously cast as rod or tube; the larger sizes are spin-cast. Common alloy compositions are CuSn5Pb5Zn5, CuSn10Pb5 and phosphor bronze CuSn10P. The phosphor bronze is the strongest and has good abrasion resistance in difficult environments, but it is the most expensive of the three alloys and requires a hardened counterface.

2.2 Aluminum Alloys

The use of steel-backed aluminum alloy bushings is confined to oil-lubricated applications, most of which are automotive, e.g., camshaft bushings, balancer shaft bearings, and bushings for the automatic transmission. The majority are lined with aluminum–tin–silicon, occasionally with the low-tin variety of aluminum–tin. Compositions are as described in Sect. 1.4.

Solid aluminum bushings, half bearings, and washers are not uncommon in the USA, the majority being made from continuously cast or spin-cast 6% tin alloy. A heat-treatable version of the alloy containing magnesium is available. Because of the high thermal expansion of aluminum, retention of fit by the bearing in a ferrous housing requires that the yield strength be increased, either by cold work or by precipitation heat treatment. Until fairly recently, solid aluminum crankshaft bearings were used in some large diesel engines; nowadays their use is confined to gas compressors.

In Germany, solid aluminum–zinc bearings of composition AlZn5Ni1Pb1Mg1Si1 were used successfully for many years as the main bearings of a light alloy, air-cooled automotive engine. The thermal expansion of the bearings matched that of the cylinder block and there was no problem with loss of fit.

Recently, two zinc–aluminum alloys containing 12 and 27% aluminum have been promoted as less expensive alternatives to bearing bronzes in general engineering applications. Their bearing properties appear to be good, but they have not yet significantly penetrated the bearing bronze market. Their expansion coefficients are higher than that of aluminum, which can give a problem with retention of fit at elevated temperatures.

2.3 Whitemetals

The steel-backed whitemetals used in the journal and thrust bearings of rotating plant are generally tin-based, SnSb8Cu3 being the most commonly used alloy. Higher strength can be obtained from alloys of more complex composition, but at the expense of loss of ductility. A zinc-containing alloy, SnZn30Cu1, is used in some marine applications. This alloy is anodic to steel, and hence affords cathodic protection to shafts in contact with seawater.

In some applications such as steam turbines, bearing temperatures approach the safe limit for whitemetal, usually accepted as 130 °C. Alternative materials are being sought with the attractive "wiping" properties of whitemetal—the ability of the bearing surface to flow plastically when the surface temperature exceeds the solidus temperature in areas of hard rubbing—but so far none has found wide acceptance.

Tin-based whitemetal bushings are firmly established in mid-range fractional-horsepower (fhp) motors, that is, with shaft diameters from 10 to 16 mm, where they are used in combination with oil-impregnated wick lubrication. Belt tension can apply high load to the bushings, and the reversing motion found in some domestic appliances compounds the problem of establishing a satisfactory oil film. The life of tin-based whitemetal in this application cannot be matched by lead-based or any other bearing alloys.

A 50-year-old US survivor in automotive applications is a lead–Babbitt-based bearing, in which a PbSn4Sb4 alloy is impregnated into a porous sintered bronze layer on a steel backing. After boring, the thickness of Babbitt above the sinter is less than 0.5 mm, and for many years this gave adequate strength for the automotive crankshaft bearing application. The material is now found only in US passenger car camshaft bushings.

2.4 Polymer-based Bearings

Metal-backed polymer-lined bearings offer a good combination of properties in sparsely lubricated applications. The high thermal expansion and low creep strength of unmodified polymers militate against their use in arduous bearing applications, as does their low thermal conductivity. These problems are largely overcome if the polymer takes the form of a thin layer firmly bonded to a metal backing. Load-carrying capacity, dimensional stability, and retention of fit become good, and conventional bearing clearances may be used. The advantages which such bearings offer over bearing alloys are superior compatibility, lower friction, and an ability to make better use of limited lubrication.

In this section, an artificial distinction is made between materials developed for unlubricated applications and those which require some degree of oil or grease lubrication. The former class, in which the primary polymer is polytetrafluoroethylene (PTFE), is dealt with in a later section, even though they have found some important lubrication applications.

Of the materials intended for use with oil or grease lubrication, a thermoplastic polymer, generally acetal copolymer, is keyed to a steel backing via a porous sintered bronze interlayer. The bronze layer is 0.25–0.35 mm thick, and there is about the same thickness of polymer above the bronze. Bushings, half bearings, and thrust washers may be formed by conventional means.

With grease lubrication, longer relubrication intervals are possible than are needed for bronze alloy bushings and this is of benefit in those applications—agricultural equipment, for example—where relubrication does not get the attention it should. The lubricated friction of polyacetal is less than that of bronze, and the use of the polymer-lined composite for kingpin bushings in a truck axle significantly reduces the steering effort.

Limitations of the acetal copolymer lining are its relatively low softening point and its susceptibility to attack by acidic oil. Composites based on more stable, high-temperature engineering polymers have therefore been developed and are now well established. One such product is based on polyetheretherketone (PEEK) and incorporates PTFE and other fillers; another is based on polyphenylene sulfide (PPS) and incorporates PTFE and Aramid powder.

Another type of polymer-based composite used in bearings is thermosetting phenolic or polyester resin, reinforced with randomly oriented asbestos fiber, or with cotton or asbestos cloth. The strength of these composites, particularly those incorporating cotton cloth, is high, and they perform well in heavily loaded applications such as rolling-mill bearings. Water makes a surprisingly good lubricant for these composites, which accounts for their use as marine stern tube bearings (carrying the propeller shaft) and rudder pintle bushes.

2.5 Oil-Impregnated Porous Bronze

Where a small-bore bushing is required, of bore diameter less than about 10 mm, an oil-impregnated porous bronze is often the most economic choice. Injection-molded nylon or acetal bushings may be cheaper, but they have a more limited load and speed capability.

Powder metallurgy techniques enable plain bushings, flanged bushings, and spherically backed bushings (when a self-aligning feature is required) to be readily produced. The third form is commonly found in the smaller fractional-horsepower motors. While there is no limitation technically to the size of bushing, the metal cost makes the product uncompetitive with steel-backed bushings in the larger sizes. Iron may be used instead of bronze, but the performance characteristics are not as good.

Compaction is controlled so as to leave a porosity of 10–25%, depending on the application. The pores are interconnected, and in service the oil circulates within the wall of the bushing. Because lubrication is hydrodynamic, oil-impregnated porous bronze bushings give their best performance when the operating conditions favor the formation of a hydrodynamic film, that is, continuous shaft rotation in one direction. The load-carrying capacity is less when the shaft oscillates, when there is frequent stopping and starting, or when the rotational speed is low.

3. Bearing Materials not Requiring Lubrication

3.1 Advantages of Dry Bearings

Many different types of bearing composite have been developed for operation without an externally applied lubricant. None is capable of operating at as high a combination of load and rubbing speed as a well-lubricated bearing—friction is higher, heat generation is greater, and there is no flow of oil to take away the frictional heat. There are, however, many advantages to dry bearings which account for their increasing popularity. The cost advantage resulting from the elimination of a lubricating system or of the labor of relubrication is a major consideration. Lubrication points are often inaccessible, and with a dry bearing the human factor in relubrication is eliminated.

There are also important technical considerations in many applications which favor the choice of a dry bearing:

(i) product contamination by the lubricant;

(ii) accumulation of airborne contaminant collected by the lubricant;

(iii) extremes of temperature—above the 150 °C maximum or below the −30 °C minimum at which conventional lubricants can be used;

(iv) absence of atmospheric pressure, resulting in evaporation of the lubricant (e.g., in space);

(v) motion that does not favor hydrodynamic lubrication—high loads and low speeds, frequent stop starts or oscillating motion;

(vi) flat rubbing surfaces and axial sliding in bushings (i.e., no hydrodynamic wedge of oil); and

(vii) operation in low-viscosity liquids (i.e., little load-carrying capacity of the hydrodynamic wedge).

Dry bearings based on polytetrafluoroethylene have additional advantages related to the frictional properties of PTFE: the low value of the friction coefficient at high loads and low rubbing speeds, and freedom from stick-slip (intermittent motion, resulting from a higher static than dynamic coefficient of friction). These advantages lead to the use of PTFE-based bearings in both lubricated and unlubricated applications.

3.2 Steel-backed PTFE–Lead–Bronze

The most popular dry bearing material based on PTFE is a steel-backed composite with a 0.25 mm lining of porous bronze into which a mixture of PTFE and lead is impregnated. A high degree of dry wear resistance derives from the presence in the rubbing surface of the three components.

The concept of *PV*—the product of load per unit area of the bearing and linear rubbing speed—is a useful way of describing the severity of a dry bearing application. Wear rate varies inversely with *PV*, and for several, though not all, dry bearing materials the wear rate is roughly the same for a given value of *PV* whether the product of high load and low speed or low load and high speed. Other factors such as temperature, the material and roughness of the counterface, and the type of motion must be taken into account, but with a knowledge of all these factors it is possible to predict the wear rate of many dry bearing materials with reasonable accuracy.

The steel-backed PTFE–lead–bronze composite has a very high degree of dry wear resistance confined to a thin surface layer, 0.025–0.050 mm thick. At *PV* of $0.5\,\mathrm{MNm^{-2}ms^{-1}}$, wear through this layer takes place at about $0.025\,\mu\mathrm{m\,h^{-1}}$. The usable temperature range is wide, from 200 °C below to 200 °C above zero, although at the higher temperature there is an increase in wear rate by a factor of ten or more. The friction coefficient is 0.10–0.25 under most sliding conditions, but falls to <0.05 under high-load, low-speed conditions (loads $>10\,\mathrm{MNm^{-2}}$, speed $<0.1\,\mathrm{ms^{-1}}$).

A lead-free version of the product is available, in which the lead is replaced by molybdenum disulfide (MoS_2). Dry wear performance is good, but not equivalent to the lead-containing product.

Applications for both products have been found in a very wide range of industries: in agricultural machinery, domestic appliances, textile machinery, machine tools, aircraft and aerospace, and civil engineering.

In two important applications for this type of composite the bushings are in fact oil-lubricated, and it was their frictional characteristics which led to their adoption. One is in the Macpherson strut, common in the suspension of front-wheel-drive cars. The other is in high-pressure versions of the hydraulic gear pump, where loads are high and thin oil films result in unacceptably high friction in a bronze bushing.

The strut application has led to the development of modified versions of the product, aimed at better satisfying the conflicting demands for low friction, high wear resistance, and resistance to cavitation erosion. One successful product incorporates fibrillated Aramid fibers in the PTFE: in another, the lead is replaced by a combination of boron nitride, molybdenum disulfide, and other fillers.

The BN/MoS_2 combination has proved successful also in unlubricated applications, and offers a performance equivalent to PTFE–lead impregnated bronze but without the environmental concerns associated with lead.

3.3 PTFE-impregnated Bronze Gauze

A product competing with the foregoing is based on a tin–bronze woven gauze impregnated with PTFE and fillers. It may be used as a floating insert or bonded to a steel backing and formed into bushings. High-volume automotive applications have been found in door hinges, pedal bushings, and gearshift linkages.

3.4 Solid Plastics

One of the advantages of the bearing construction described in Sect. 3.2 is that frictional heat is readily removed through the bearing wall, both the porous bronze lining and the steel backing being reasonably thermally conductive. Solid plastic bearings are at a disadvantage in unlubricated applications, in that frictional heat must be removed through the mating surface, which may or may not be a good heat sink. For this reason, and for the reasons of dimensional instability mentioned in Sect. 2.4, solid plastic dry bearings are restricted to applications in which *PV* does not exceed $0.2\,\mathrm{MNm^{-2}\,ms^{-1}}$.

PTFE is the polymer on which most dry bearings in the solid plastic category are based. The PTFE may be the matrix in which such fillers as graphite or glass fiber are incorporated, or it may itself be a filler in a thermoplastic matrix such as polyacetal. In one well-established product of the latter type, the PTFE is present as randomly oriented fibers in an acetal copolymer matrix.

Unmodified PTFE is rarely used as a dry bearing material, on account of its high wear rate and low creep strength. Interesting exceptions are, however, found in civil engineering: the bearing pads supporting bridge spans and flyovers often consist of a slab of PTFE confined in a shallow cast iron well, to prevent extrusion under the very heavy bearing loads. PTFE offers a friction coefficient as low as 0.02 under these conditions, and permits thermal expansion and contraction of the span without imposing undue stresses on the structure.

The low creep strength of PTFE may be improved by the orientation of the molecules which occurs during the drawing of fibers. The cloth woven from PTFE fibers has a high load-carrying capacity, and is found as the lining of bushings in earthmoving equipment and in aircraft control linkages. The backing is usually a thermosetting plastic such as phenolic or epoxy resin reinforced with glass fiber.

Non-PTFE composites incorporate alternative solid lubricants such as graphite or molybdenum disulfide or, in some cases, micropockets of lubricating oil. At the lower end of the cost spectrum are found the injection-molded nylons and acetals, which are both intrinsically cheap and can be processed by low-cost methods. The incorporation of solid lubricant or oil improves the friction coefficient and wear resistance, and provides satisfactory performance in many lightly loaded applications. The maximum PV at which acceptable wear rates are obtained is $0.1\,\mathrm{MNm^{-2}ms^{-1}}$. In the oil-containing version, the oil pockets are not interconnected and the performance does not compare with that of oil-impregnated bronze.

At the upper end of the cost spectrum are expensive, high-temperature engineering polymers such as polyimide. Graphite-filled polyimide provides good temperature resistance and a low wear rate in many difficult applications. The wear rate is relatively unaffected by increasing temperature, and acceptable performance is obtained at PV values up to and sometimes over $0.20\,\mathrm{MNm^{-2}ms^{-1}}$.

Intermediate in cost are thermoplastics such as polysulfone and polyphenylene sulfide reinforced with randomly oriented carbon fiber. Quantitative data on the bearing performance of these composites are not yet available.

3.5 Graphited Bronze and Carbon

Bronze and steel-backed bronze bearings incorporating pockets of a graphite-based solid lubricant in the bearing surface are used where some enhancement of wear resistance is required under sparse lubrication conditions. For operation without lubricant, a much finer distribution of graphite throughout the bronze matrix is required. Composites of this type are produced by powder metallurgy techniques; copper, tin, and graphite powders are blended, compacted, and sintered. The most common composition contains 8 wt.% (26 vol.%) of graphite. Up to 40 vol.% of graphite can be incorporated, though at some considerable sacrifice of strength and ductility. The degree of dry wear resistance is not as high as that obtainable from the PTFE-based composites, but the graphited bronzes may be used at temperatures well above those that PTFE can withstand. They are found in hostile environments where a certain amount of wear can be tolerated, and also in some underwater applications.

Carbon bearings also have good temperature resistance, and may be used in lightly loaded applications at temperatures as high as 500 °C. They have excellent corrosion resistance and are used in pumps handling corrosive liquids. Their thermal expansion is low, and special measures have to be taken to retain fit in the housing. The counterface is usually hard and corrosion resistant.

3.6 Ceramics

Ceramics such as alumina and silicon nitride, and cemented carbides such as cobalt-bonded titanium and tungsten carbides, are used successfully against hard counterfaces in certain specialized applications. An excellent surface finish may be obtained on these hard materials; for this reason they are used in gas-lubricated bearings, in high-precision machine tools, and as the pivot bearings in horological and instrument applications. They are also found as very high-temperature bearings, and in nuclear reactors on account of their resistance to radiation.

Alumina is used as the water-lubricated thrust bearing in domestic hot water circulating pumps. Cemented carbides are used as the bearings of industrial pumps handling corrosive liquids. Their high abrasion resistance enables them to provide satisfactory performance when the medium in which they are operating is heavily contaminated by abrasive particles, such as in mining machinery.

See also: Automotive Body Trim, Coatings, and Paint Systems#; Automotive Engine Materials#; Wear and Lubrication#

Bibliography

Braithwaite E R (ed.) 1967 *Lubrication and Lubricants*. Elsevier, Amsterdam

Brown W P, Galand H, Kingsbury G R, Scott M 1996 Aluminium bi-metal bearing developments for automotive and medium speed diesel applications. *Ind. Lubrication Tribol.* **48** (3)

Kumada Y, Hashizume K, Kimura Y 1995 Microgrooved bearings—new concept engine bearings of higher performance. *Proc. Int. Tribology Conf.* Japanese Society of Tribologists, Yokohama, Japan

Morris J A 1967 Metallic bearing materials. In: Braithwaite E R (ed.) *Lubrication and Lubricants*. Elsevier, Amsterdam, Chap. 7, pp. 310–76

Pratt G C 1967 Plastic-based bearings. In: Braithwaite E R (ed.) *Lubrication and Lubricants*. Elsevier, Amsterdam, Chap. 8, pp. 377–426

Pratt G C 1973 Materials for plain bearings. *Metall. Rev.* **18**, 62–88

Pratt G C 1977 The wear properties of polymer composites. In: Richardson M O W (ed.) 1977 *Polymer Engineering Composites*. Applied Science, London, Chap. 5, pp. 237–61

Whitney W J An advanced aluminium–tin–silicon engine bearing alloy and its performance characteristics. Society of Automotive Engineers Technical Paper 950953

G. C. Pratt
Isle of Harris, UK

Biological Materials: Synthesis, Structure, and Properties

The microstructure and properties of biological tissues are of interest because they influence the morphology and behavior of each species. Biological tissues are also of interest in that they can suggest new designs for synthetic materials and new ways of using synthetic materials to achieve an engineering goal. In surgery, we have the more difficult problem of developing synthetic materials that will work in tandem with animal tissues. The following article contrasts synthetic and biological approaches to processing, to material structure and to the use of materials in design. It also provides links to more detailed descriptions of individual types of plant and animal materials in this encyclopedia.

A key observation concerning biological materials is that they are all composites. At any scale above 10 μm there are no uniform structures in biology. This makes the concept of a "biological material" rather uncertain because structures and properties change with position in the body, with the individual plant or animal and with time. Thus we can consider "cartilage" to be a material, but then distinguish the structures of hyaline, fibrous, and elastic cartilage. We then subdivide each of these according to their site, such as articular cartilage and costal cartilage. Also, in contrast to assembled machines, sharp boundaries between structural materials are rare so it is quite difficult to define where one material stops and the next one starts. For these reasons it is also relatively difficult to characterize tissues and compare their properties with those of plastics, ceramics or metals.

1. Biological and Synthetic Processing

In practical engineering, where cost is very important, there is little sense in designing a part without a very good idea of how it will actually be made. Materials selections are then made once the manufacturing route and basic design have been decided. The same must be true in biology; many interesting structural features may be firstly a consequence of the growth process and only secondarily a source of improved properties.

Small numbers of synthetic parts can be made by subtractive processes such as machining but the process is slow and wasteful. Molding is generally cheaper and faster. Hot liquid material is injected into a hollow tool and allowed to solidify. Extrusion, forging, rolling, and other hot processes can be viewed as variants on molding. The sintering of ceramics is a separate case; compacted powder is heated until surface energy drives the slow shrinkage and densification to a solid. Chemical processing is not common in the synthetic world. Thermosetting resins such as epoxies are solidified by chemical reaction (see *Thermosets: Epoxies and Polyesters*[#]) and there is some exotic reactive processing of ceramics. Generally, chemistry has the problem that small changes in starting conditions or purity can cause big changes in the reaction kinetics and so reliability is poor.

In biology, all processing is essentially chemical. Soluble reagents are fed to a site where they combine to make a solid plus dissolved by-products. In the case of epoxies, mentioned above, there is little volume change between the liquid and solid states. In forming solids from solution, there is also a massive shrinkage to be accommodated. The biological equivalent of molding can be seen in the formation of isolated particles within an envelope of lipid membrane. Examples include the formation of silica in sponge spicules (Simpson and Volcani 1981) and diatoms (Li and Volcani 1984) and calcite in coccolith (Young *et al.* 1992) skeletons. The wall of the "mold" is now permeable to allow reagents in and soluble products out. This approach is suitable for isolated particles, such as the magnetic particles in magnetotactic bacteria (Frankel *et al.* 1998) and for the wall of a single celled diatom. The particles may also be later assembled into a simple framework, as in the sponges. However, this method does not lend itself to strong, dense, load-bearing structures suitable for large plants and animals.

A problem with building a large solid object by a chemical precipitation is to avoid surrounding and entombing the cellular machinery that provides the reagents and removes the products. The obvious solution is to build the solid layer-by-layer, such that a layer of cells provides material to add to the surface of the growing solid and retreats ahead of it. The layer approach is apparent in the growth rings of trees. Bone, tooth and shell form in the same way. In wood and skin, it is the cells themselves that become the structural unit by depositing solid cellulose or keratin in or on the cell. Each new layer starts as a new layer of cells.

In bone and tooth the deposition is external to the cells (see *Bone Mineralization*[#]). In the case of bone growth, the cell layer continuously forms new collagenous matrix material. This then mineralizes and converts to hard bone as new matrix is deposited over it. The control system thus promotes crystal growth in layers several microns from the cell surface while not mineralizing the freshly formed matrix. Many proteins are known to be associated with control of this process but it is not yet clear how they work together.

One natural consequence of this layerwise growth process is that layers of different material can readily be formed within a single solid. As will be seen below, biology makes extensive use of layered structures to add toughness to strong materials. Recent developments in ceramics have also focused on the use of layered structures to add toughness. New methods of freeform fabrication (see *Freeform Fabrication*#) should ultimately allow the production of complex layered structures resembling those of biological materials.

A result of the chemical precipitation is that it is relatively difficult to make dense structures. The large volume change as a soluble polymer assembles into fibers or as a mineral precipitates will lead to highly porous structures unless the deposition rate is extremely slow. Any structure forming in a diffusion field will tend to grow towards the source rather than filling in gaps in a layer. Very slow growth will allow equilibration, which will favor dense structures. One solution to this is to lay down a mesh of strong fibers, fill in the pores with a softer matrix and then arrange for this to slowly expel water and harden chemically. This kind of process is seen in wood and in cuticle (see *Cuticle*#) and very similar issues occur in the formation of carbon–carbon composites (see *Carbon–Carbon Composites**). For such reasons, biological materials must be composites.

2. Structural Biological Materials

In their survey of biological materials, Wainwright *et al.* (1986) make a division into tensile materials, rigid materials and pliant materials. Following a more conventional materials division, we will discuss polymers, ceramics and mineralized polymers, and gels. However, such boundaries are even less clear in biology than in the synthetic world.

Two processing-induced limitations should be recognized for structural polymers in biology. Firstly, they are all formed from aqueous solution and so are all very sensitive to plasticization by moisture. In most cases, the properties of the dry material have little relevance because they will only occur in a dead organism. The plasticizing effect of water on biological polymers can be regarded as parallel to the softening effect of increased temperature on synthetic amorphous polymers. As temperature will take a hard polymer through the glass transition into a rubbery state, so increased water content will convert amorphous proteins from glass to rubbery. There is little sense in measuring the mechanical properties of biological materials without defining the water content.

Secondly, the growth process allows a variety of routes to the formation of fibers but there are no simple ways to form a dense isotropic plastic. Thus, even essentially isotropic materials will have a fibrous composite microstructure. Weiner (Weiner *et al.* 2000) has recently argued that many biological structures can be viewed as a search for isotropic properties, or at least orthotropic properties (strong in two dimensions), from fibrous materials. This would reflect the unpredictability of stresses encountered by a structure in a dynamic environment.

Tables 1 and 2 compare the mechanical properties of synthetic and biological materials.

3. Structural Proteins

Many insects and arachnids make extensive use of silk protein fibers for a range of purposes (see *Natural Protein Fibers**). Spiders typically produce five different silks for the radial and spiral parts of the web, for a sticky web coating, for wrapping prey and for the dragline. The property variations come from differences in the amino acid sequence of the polymer. Silks are stored in a gland as an aqueous solution, at least some of the time, in a lyotropic liquid crystalline state. For the strong silks, shear at the spinneret leads to a change in conformation to a very highly oriented and stiff fiber. The combination of strength, stiffness, and toughness shown by spider dragline silks have led to efforts to characterize, clone, and produce a synthetic version of this material (see *Silk Produced by Engineered Bacteria*#).

Bulk crystalline polymers, such as nylon or polyethylene, have an amorphous fraction of 30–50% that arises from the inability of entangled polymer chains to become completely ordered. This crystallinity can be controlled to some extent by processing and can be reduced by copolymerization to introduce random irregularity into the chain. Fiber structures are less easy to resolve but behave similarly. In many

Table 1
Properties of natural and synthetic ceramics

	Modulus (GPa)	Strength (MPa)	K_{1c} (MPam$^{1/2}$)	G_c (Jm^{-2})
Alumina	350	100–1000	3	7
Fused silica	72			9
Nacre, wet	64	130		1240
Nacre, dry	73	167		464
Enamel	45	76		13–200
Nephrite (jade)	222		2.9–5.5	269–619
Horneblende		151	7–16	49
Jadeite	205	205		94–120

Table 2
Properties of synthetic and biological composites

	Volume fraction (%)	Tensile modulus (GPa)	Tensile strength (MPa)	Strain to break (%)	Work of fracture (Jm^{-2})
Continuous fiber PEEK/AS4, parallel	61	140	2200	1.5	1900
Continuous fiber PEEK/AS4, perpendicular	61	8.3	73		
Sheet molding compound (resin + filler + fiber)	50, 15–30% fiber	15	86–230	1.3–1.7	
Polybutyleneterephthalate/ glass beads	25	4.9	95		
Polyethyleneterephthalate + short glass fiber	35	20	165	1	3200
Polyethylene-hydroxyapatite	40	5	22		1800
Polyethylene-hydroxyapatite	50	9	26		1000
Polyethyleneterephthalate	0	3.3	60	275	7300
E glass	100	70	3000		
Macadamia nut shell	0		25–80		100–1000
Insect cuticle	0	6–10	95		1500
Antler	31	7.7	179		6200
Bone (bovine femur)	41	20	220	10	1700
Dentine	48	12	250		550
Whale bulla	66	30	33		200
Collagen (tendon)	0	3	100		
Hydroxyapatite	100	110	100		

silks the structure seems to be blocky, with irregular sections spaced along the chain to define noncrystallizable sections. We are not yet able to resolve the role of these irregular sections in stabilizing the liquid crystalline state and in defining the final structure and properties of the fiber. In all biological polymers there is a degree of control over the molecular structure that may be very important for properties and cannot be duplicated synthetically.

Collagen is the structural material of skeletal animals and has properties that are quite inferior to those of cellulose, chitin or silk. The key to its use seems to lie in the versatile processability. Soluble procollagen is formed in the cell and exported into the growing tissue. An enzyme cleaves a bulky end section from the molecule to form tropocollagen that organizes into a triple helical structure. These triple helices self-assemble into collagen fibrils that make up the bulk of tensile structures such as tendon and ligament. As it ages, the collagen becomes cross-linked, which increases the stiffness and strength but reduces the toughness. Baer *et al.* (1991) have discussed the structure and properties of collagen in tendon and ligament.

Keratin can be seen as a biological answer to the need for a tough plastic equivalent to nylon. As hair and fur, it is a fiber. As epidermis it is a film and as hoof or horn, it is tough solid. The structure contains fibrils, which are built from three-stranded ropes of alpha-helical chains in a coiled-coil arrangement. The fibrils are embedded in a cross-linked matrix of amorphous protein that is heavily cross-linked by cystines (sulfur–sulfur links). The high sulfur content is a defining characteristic of members of the keratin family and gives burning hair its characteristic smell. The fibrils confer great toughness by converting to a more extended beta-sheet structure under stress. Feather keratin is a beta-sheet structure presumably for higher stiffness. Amphibian skin is also a beta-sheet, again presumably to limit water swelling. The best-studied example of keratin is wool but there is also recent work on hoof (see *Keratin*$^{\#}$).

4. Structural Polysaccharides

Cuticle is a composite of chitin fibers embedded in a cross-linked protein matrix (see *Cuticle*$^{\#}$). Chitin is a polysaccharide, similar to cellulose but with acetylamino substitution. Metabolically, polysaccharides should be cheaper than protein since nitrogen has a much lower abundance in the biosphere than carbon, hydrogen and oxygen. It is not clear what improvement in properties or processing led the insects to select chitin in the place of cellulose as their main structural material. The layered structures of cuticle, with various sequences of fiber orientation, do strongly resemble the layered structures of carbon-fiber composite laminates. Given that both insects and military aircraft are lightweight, roughly cylindrical systems, the resemblance cannot be accidental. Gunderson and Schiavone (1995) have discussed how insects adopt unbalanced or asymmetric lay-ups that would not be used in synthetic composites. Some of the patterns also involve thick layers oriented along the major structural axes, with many fine layers forming the rotation from one direction to the next, apparently to delocalize shear stresses that would cause delamination.

Many layered biological systems resemble cholesteric liquid crystals in the rotation of orientation from layer to layer. This has raised interest in whether they actually are liquid crystals, in that the rotation forms spontaneously as a result of interactions between fibers in successive layers (Giraud-Guille 1996). This would occur in a fluid state, which is subsequently embedded in a hard matrix. The core question is really whether the rotation pattern is directly controlled by some form of oriented extrusion during the deposition process or is controlled through the surface chemistry of fibrils deposited in successive layers.

The mechanical properties of chitin are hard to define because large oriented samples are unavailable. In cellulose, many plants contain bast fibers with very highly aligned polymer that allows us to calculate the stiffness of the polymer and so analyze the properties of wood and other plant materials. The stiffness of cellulose in plant fibers, 40 GPa wet and 100 GPa dry, is comparable to the stiffest synthetic fibers, such as Kevlar. The polymer chains are wholly aligned with the fiber axis.

Plant cells are hollow tubes with spirally wound fibrils. Successive layers go in opposite senses, making a criss-cross pattern. The outer and inner layers may be wound at different angles to the rest. This is very like the winding of fibers on a composite pressure vessel, such as a pressurized gas tank. The wood composite structure is designed to retain internal pressure or longitudinal compression with very high energy absorption on fracture as the windings collapse inwards.

Plants again have the problem of how to manufacture the polymer without it enveloping and choking the production site. It is laid down by “track-laying” system, which seems to carry out the polymerization at the cell surface on a moving organelle and which draws glucose for polymerization through the membrane from inside.

In wood, hemicelluloses and lignin act as a matrix bonding the cellulose fibers into a composite structure. It is a puzzle that the stiffest, strongest and cheapest of the biological polymers is not used at all in animals.

5. *Mineralized Tissues*

In the synthetic world, some applications involve predominantly tensile loading; the wall of a pressure vessel is one example. Much more commonly, parts will be loaded in compression or bending. While strong fibers in the form of a rope can provide excellent resistance to tension, they are of little use in compression. Large animals with an external or internal skeleton will need stiff materials that are more isotropic in their properties and so can withstand the varying stresses that come from moving around and colliding with other objects. Mineralized composites provide improved compressive properties over purely polymer structures. Compared to stiff polymers, minerals can also be formed at a lower metabolic cost for a given level of stiffness, but do increase the overall weight.

A wide range of minerals is found in microbes, plants and animals but silica, calcium carbonate and hydroxyapatite are the most important.

6. *Silica*

Silica occurs as spicules, short reinforcing rods, in sponges. The spicules are typically 10 μm in diameter and 100 μm long and may be simple rods or complex branched structures. They apparently form by aggregation of silica nanoparticles onto a thread of polysaccharide or protein within a vesicle, a membrane-enclosed space inside the organism. The material is amorphous, highly hydrated and not fully dense. There is one example of a deep-sea sponge, *Monoraphis*, that is attached to the seabed by a large, one meter long, silica rod. The structure of this material is a series of concentric layers where the weaker interlayer regions may enhance the toughness of the structure, like a laminated glass windshield (Levi *et al.* 1989).

Siliceous diatoms are single-celled prokaryotes, about 100 microns in diameter, enclosed by a porous silica shell, in the form of two overlapping dishes, like a petri dish. The organism multiplies by dividing into two disc-shaped cells and then forming two new dishes back-to-back. The cells then separate, each with one new half shell and one old one.

Recent work on the proteins involved in silica mineralization has started to clarify the details of silica deposition in diatoms (Kroger *et al.* 1999) and sponges (Morse 1999).

Plants often contain silica as a reinforcing material. The structure of bamboo has been much studied in this context.

Sea water and soil water contain low levels of dissolved silica as silicic acid. The cell probably binds and transports this as a complex with catechols (dihydroxybenzene) and then converts this to silica nanoparticles.

7. Carbonate

Calcium carbonate is widespread as a protective shell in marine animals including single-celled coccoliths, coral, gastropods and bivalves (Lowenstam and Weiner 1989). It also occurs as a reinforcement in the cuticle of crustaceans such as crabs. The use of calcium carbonate for shells, rather than silica, may reflect the greater control of structure available through calcium-binding proteins and through control of crystal morphology with nucleators and growth inhibitors. Silica does occur as a component of many mollusk teeth, as do various iron oxides.

There has been much work on the formation, structure, and properties of mollusk shell (see *Shell: Properties*$^{\#}$). This has been partly driven by a desire to understand why these structures are much tougher than the natural mineral. This insight could allow us to make synthetic ceramics much more useful. Most shells are constructed of calcite and aragonite in various arrangements with small amounts of protein, up to about 5% by weight. Higher protein levels would be expected to increase the toughness of the shell but at the expense of reduced stiffness. It can be assumed that the particular shell structures are adapted to the particular lifestyle of the animal and the stresses encountered. For instance, it might be expected that a swimming bivalve, such as a scallop, would have a stiffer, lighter shell structure. One of the most studied shell structures is nacre, or mother-of-pearl, which is a layered structure of aragonite plates, 0.5 μm thick by several microns wide. Between each layer of plates is a thin protein sheet that is responsible for the nucleation of the aragonite layers and acts as a crack stopper. This gives wet nacre a high fracture energy, though this toughness is lost on drying or freezing, Fig. 1.

The characteristic fracture of these tough natural ceramics is ‘‘graceful failure,’’ shown by an extended sequence of crack formation and arrest that gives the saw-tooth stress-strain curve of Fig. 1. Conch shell shows a similar fracture curve but this arises from crack deflection within the complex lamella architecture rather than from protein layers (Kamat *et al.* 2000). In sea urchin spines there is much protein embedded as microspheres within the mineral crystals but it is not known whether this really contributes to toughness.

While the details of shell formation are not understood, it is clear the organic content contains proteins capable of selectively nucleating calcite or aragonite in a specific orientation and proteins which inhibit growth of specific crystal faces. These work in concert with the supply of calcium and carbonate from the mantle tissue. The mantle adds new shell at the existing edge by extruding over the lip to extend the outer and inner faces.

The vaterite form of calcium carbonate often occurs in laboratory crystallization and is occasionally found in shells. There is increasing evidence for precipitation of amorphous, hydrated calcium carbonate as a metastable precursor for the crystal (Raz *et al.* 2000). In some cases, including lobster cuticle and a sponge, stable amorphous calcium carbonate is found (Beniash *et al.* 1999).

8. Hydroxyapatite

The human body is supersaturated for both calcium carbonate and for hydroxyapatite, $Ca_{10}(OH)_2(PO_4)_6$, the mineral of bone and tooth. Since the environmental availability of phosphorus is limited, its use as

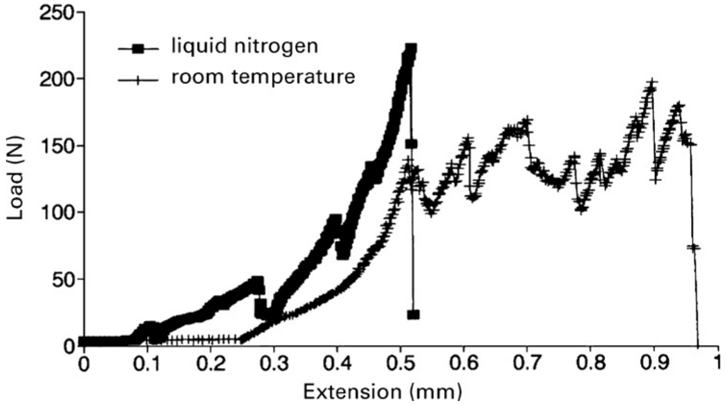


Figure 1
Fracture of a nacreous shell in bending at room temperature and at liquid nitrogen temperature. Freezing the protein converts the tough material into a brittle material like a conventional ceramic.

a reinforcement seems peculiar. Possibly bone acts as phosphorus reservoir. Bone is essentially a polymer-matrix composite reinforced with ribbons of hydroxyapatite. The structure and properties of bone are discussed elsewhere (see *Bone and Natural Composites: Properties**; *Bone Mineralization*#). A key current question is the way the lamella structure gives rise to a combination of high stiffness, high strength and high toughness. Similar synthetic composites tend to break at relatively low strain and so have low toughness.

Dentine has a similar structure to bone but with significant differences in the mechanism of mineralization. Tooth enamel is almost wholly mineral with a fibrous structure that forms under the control of a completely different set of mineralization proteins, including amelogenin.

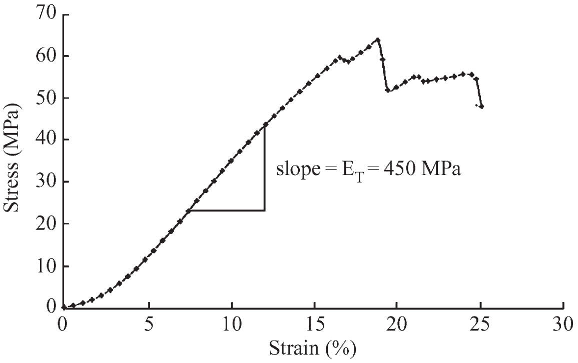


Figure 2
Stress-strain curve for human Achilles tendon, showing the J-shaped response. (Reproduced by permission of IOS Press from *Biomed. Mater. Eng.*, 1997, **7**, 231–44)

9. Soft Materials

Many of the materials in the human body are water-swollen polymer composites that are soft but strong. Dry proteins are stiff and brittle but a hydrophilic composition leads to increased water content, which acts in the same way as increased temperature in a synthetic polymer and moves the material through the glass transition into the rubbery region.

Elastin is such a rubbery biological polymer, though this only occurs pure, more or less, in the *ligamentum nuchae* that supports the head of grazing animals. Other biological rubbers are abductin in the hinge of bivalve mollusks and resilin, which drives the elastic rebound of insect wings.

Many soft biological materials are more like textiles with a three-dimensional arrangement of stiff fibers embedded in a hydrogel matrix. Various weaving patterns allow the material to extend greatly in directions perpendicular to the local fiber axis, while remaining stiff parallel to the fibers. Early extension occurs as a result of fiber rotation and the material then becomes stiffer when the fibers become oriented. The dermis (skin), arterial and intestinal wall are such combinations of stiff collagen fibers, rubber elastin fibers, water and proteoglycan (protein with polysaccharide side chains) gel. Mesoglea, in the body wall of the sea anemone, is a composite of collagen and polysaccharide gel. These materials also contain muscle fibers and some have reversible cross-linking mechanisms to allow the fibers to slide. This permits a combination of short-term stiffness with very large slow extensions.

One characteristic that most of these soft materials share with textiles is a J-shaped stress–strain curve (Fig. 2). The initial response is with a low elastic modulus but the material becomes much stiffer at high strains. This form of response allows a combination of softness and flexibility with high strength. In the arterial wall, for instance, this stress–strain relation allows the artery to expand and contract under normal pressure pulses while resisting balloon expansion if the pressure becomes locally high.

Cartilage contains about 50 vol.% collagen fibers in a proteoglycan and water matrix. The highly-sulfated polysaccharide component provides a high local ionic strength which causes the mesh to swell under the fibers provide an opposing stress, like a balloon inflated inside a sock. This structure shows strongly time-dependent elastic behavior as water moves in or out of the structure. There are still no good medical replacements for worn articulator cartilage in arthritic joints. The structure and properties of cartilage change with site. Costal cartilage occurs in the meniscus of the knee, the nose, ear and connection of the ribs in the sternum.

10. Active Biological Materials

A robotic system requires a skeleton, sensors, actuators, microprocessor and power supply. An animal can be viewed as the same set of subsystems, but unlike a robot, they are integrated into a monolithic body. As we seek ways of integrating active functions into materials, there is more interest in the materials properties of biological sensors and actuators.

For a discussion of the structure and properties of the eye see *Biological Optical Materials*#. Current studies of sensing and the nervous system are in the realm of membrane biology, which does not yet really allow a materials approach. The question of biomimetic sensing systems is treated in *Artificial Noses and Tongues*# and *Tactile Sensors*#.

The materials properties of muscle are discussed in *Natural Muscle*#. In robotic terms this is a chemically powered actuator with a linear contraction of around 50% and excellent force control. On the other hand, the speed and peak stress are relatively low. There is a great need for an equivalent synthetic system, preferably electrically powered (Liu and Calvert 2000).

11. Biological and Synthetic Design

A number of differences between biological and synthetic materials have been discussed above. In the synthetic world, the choice of material is often made in the middle of the design process and the design is then adjusted to the properties and processing methods for the material. In biological systems, the material structure and properties are part of the design.

Small animals, such as insects, have hard shells containing the active components. Most appliances and vehicles follow the same design. Large animals have skeletons, to carry largely compressive and bending loads, which support the active components. Large animals are much less susceptible to collision damage than most vehicles.

It may be sensible to adopt biological designs for materials and systems more in the synthetic world. However, the only real principle of evolutionary design may be that there isn't one. This is discussed further in the articles on biomimetic materials (see *Biomimetic Materials: Properties and Processing**) and design optimization (see *Mechanical Optimization in Nature, Principles of*#). In considering biological design, it should be kept in mind that the need to reproduce and grow impose constraints that do not apply to most synthetic systems. Recycling is clearly one area where we could learn from biology, but this is a property of the ecosystem rather than of the individual plant or animal.

Biomimetic materials are discussed in a separate article. Here the goal is to develop synthetic materials, which take advantage of biological composite structures, which allow machines to be built with more of the soft, flexible characteristics of animals or to build machines that function like animals. Biomedical materials are reviewed in *Biomedical and Dental Materials: Introduction*#. The need to interface synthetic materials with the body is more demanding than simply designing a synthetic material with the properties of the natural one.

See also: Biosynthesized Materials: Properties and Processing#

Bibliography

Baer E, Cassidy J J, Hiltner A 1991 Hierarchical structure of collagen composite systems: lessons from biology. *Pure Appl. Chem.* **63**, 961–73

Beniash E, Addadi L, Weiner S 1999 Cellular control over spicule formation in sea urchin embryos: a structural approach. *J. Struct. Biol.* **125**, 50–62

Frankel R, Bazylinski D, Schuler D 1998 Biomineralization of magnetic iron minerals in bacteria. *Supramol. Sci.* **5**, 383–90

Giraud-Guille M 1996 Twisted liquid crystalline supramolecular arrangements in morphogenesis. *Int. Rev. Cytol.* **166**, 59–101

Gunderson S L, Schiavone R C 1995 Microstructure of an insect cuticle and applications to advanced composites. *Biomimetics*. AIP Press, Woodbury, NY, pp. 163–98

Kamat S, Su X, Ballarini R, Heuer A H 2000 Structural basis for the fracture toughness of the shell of the conch, *Strombus gigas*. *Nature* **405**, 1036–40

Kroger N, Deutzmann R, Sumper M 1999 Polycationic peptides from diatom biosilica that direct silica nanosphere formation. *Science* **286**, 1129–32

Levi C, Barton J L, Guillemet C, Lebras E, Leheude P 1989 A remarkably strong natural glassy rod: the anchoring spicule of Monoraphis sponge. *J. Mater. Sci. Lett.* **8**, 337–9

Li C W, Volcani B E 1984 Silicification in diatom walls. *Phil. Trans. Roy. Soc. London* **B304**, 519–28

Liu Z, Calvert P 2000 Multilayer hydrogels as muscle-like actuators. *Adv. Mater.* **12**, 288–91

Lowenstam H A, Weiner S 1989 *On Biomineralization*. Oxford University Press, Oxford

Morse D 1999 Silicon biotechnology: harnessing biological silica production to construct new materials. *Trends Biotechnol.* **17**, 230–2

Raz S, Weiner S, Addadi L 2000 Formation of high-magnesian calcites via an amorphous precursor phase: possible biological implications. *Adv. Mater.* **12**, 38–40

Simpson T, Volcani B 1981 *Silicon and siliceous structures in biological systems*. Springer, Berlin

Wainwright S A, Biggs W D, Currey J D, Gosline J M 1986 *Mechanical Design in Organisms*. Princeton University Press, Princeton, NJ

Weiner S, Addadi L, Wagner H D 2000 Materials design in biology. *Mater. Sci. Eng. C* **11**, 1–8

Young J, Didymus J, Bown P, Prins B, Mann S 1992 Crystal assembly and phylogenetic evolution in heterococcoliths. *Nature* **356**, 516–8

P. Calvert
University of Arizona, Tucson, Arizona, USA

Biomimetic Materials: Properties and Processing

The concept of biomimesis has long been used in chemistry in the context of compounds with enzyme-like catalytic action. Since the mid-1980s it has been applied to materials, particularly with a view to producing ceramic and composite materials with improved toughness, analogous to shell, tooth, and bone. There was some discomfort in the materials community with the idea of mimicking biological materials in the sense of producing an indistinguishable copy.

However, mimesis more generally refers to copying some essential aspects of a thing rather than duplicating or faking it. The phrase "bioinspired materials" is also used to express the idea in more familiar terms, but is grammatically less desirable as it is a

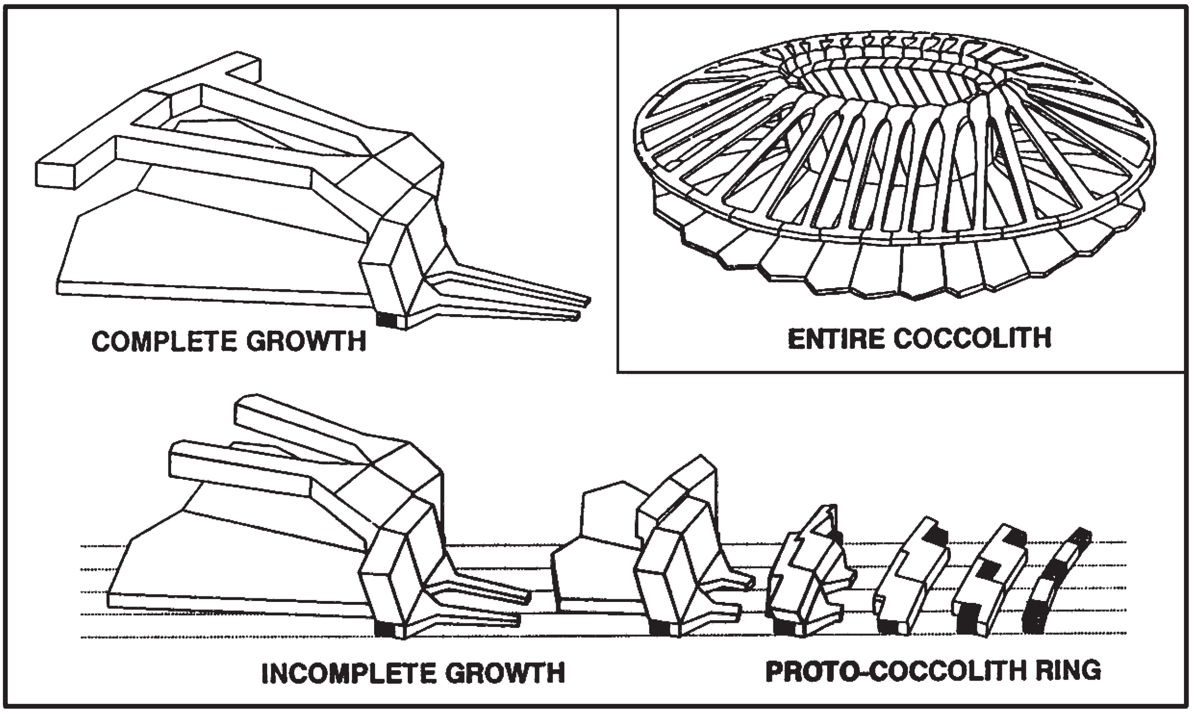


Figure 1
Assembly of a coccolith from calcite crystals (reproduced by permission of Macmillan from *Nature*, 1992, **356**, 516–8).

Greek/Latin hybrid. The field has now spread to include a group of loosely linked goals in new materials and processes. These are surveyed below.

It should be kept in mind that materials development is considerably upstream from the development of new products. It is quite typical for new materials to find their way into commercial products about 20 years after their discovery (Calvert 1994). Examples include Kevlar, high-T_c ceramic superconductors, piezoelectric polymers, and gallium arsenide. It is also typical for the first applications to be quite different from those originally proposed, and for their impact to be modest compared to that suggested during the initial excitement. A personal view of the status and prospects for biomimetic materials is given at the end of this article.

1. *Inorganic Particle Formation*

Coccolithophores—single-celled marine algae—assemble an external skeleton from single crystals of calcite with very complex shapes, such as that shown in Fig. 1. The growing crystal is surrounded by a lipid membrane that controls the growth orientation in the crystal, but it is not known exactly how this is achieved. As shown in Fig. 1, a series of overgrowths on the initial ring lead to the whole calcite assembly. Sponges and diatoms show similar close control of the shape of silica particles on the micron scale. Magnetotactic bacteria form single-domain iron oxide (magnetite) crystals with a very closely controlled size of a few nanometers, which then aggregate into magnetic chains (Fig. 2). These biological examples all involve growth within a compartment surrounded by a membrane. For sponge spicules there is an organic template on which the mineral grows. In other cases there may be specific nucleation sites on the membrane surface. This suggests synthetic approaches where inorganic particles are grown within a micromold—a predefined space.

There have been many efforts to grow particles in liposomes, i.e., spherical shells with a lipid bilayer wall (Mann *et al.* 1997). One solution is trapped when the liposome is formed, and precipitation occurs when a second reagent, often a base, diffuses in through the wall. Generally the trapped solution must be quite dilute in order to avoid destabilizing the liposome. As a result, the precipitate therefore only occupies a small part of the internal volume. A method is needed to introduce a continuous feed of both reagents through the membrane.

There have also been many efforts to grow particles in multiphase polymer systems, such as block

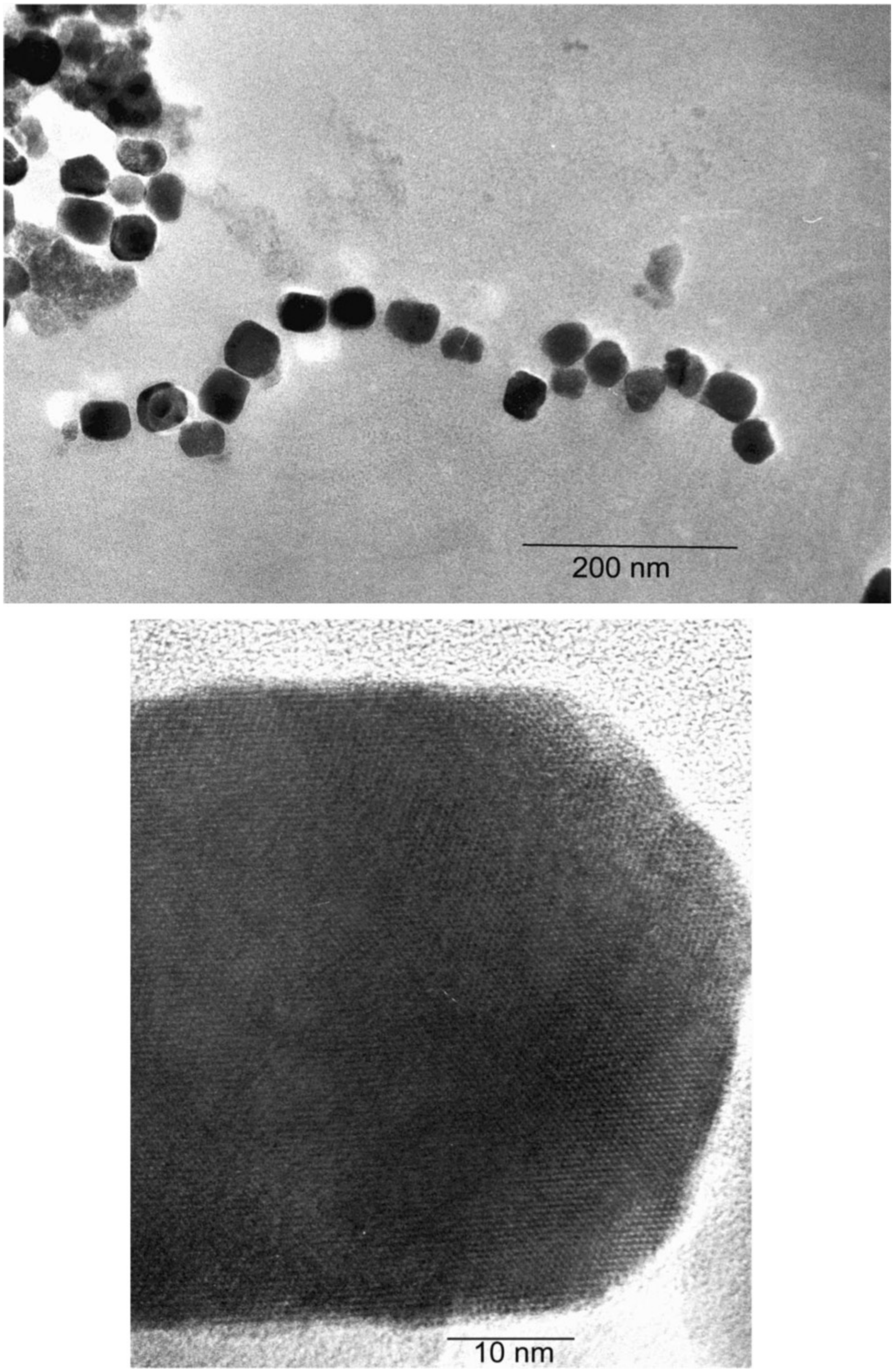


Figure 2
Magnetic particles from a magnetotactic bacterium, showing chain formation (courtesy of Professor S. Seraphin, University of Arizona).

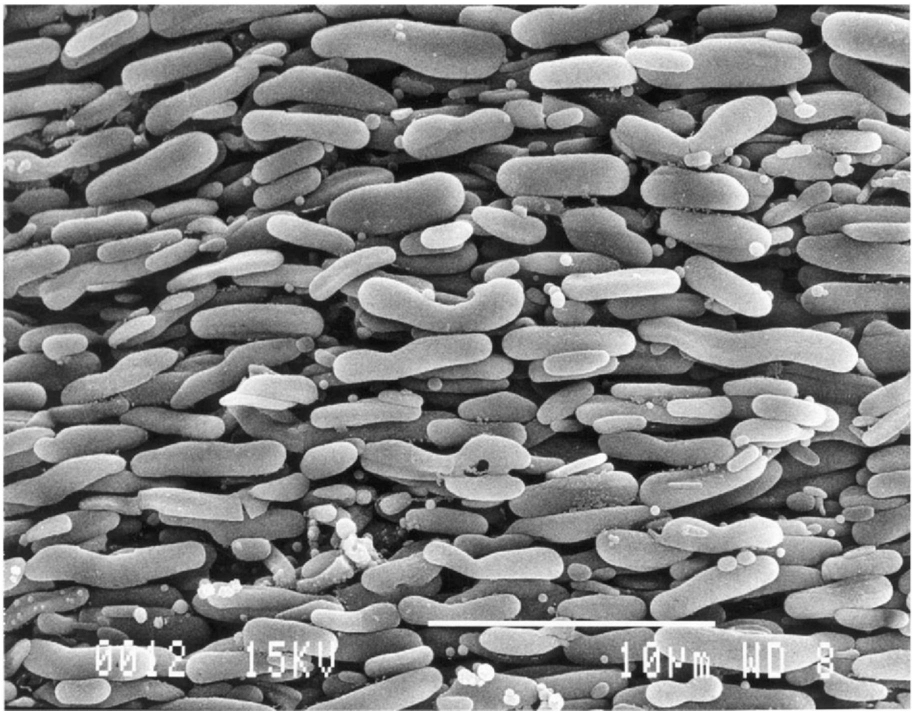


Figure 3
Elongated titania particles formed by impregnation of a stretched two-phase polymer with titanium alkoxide, followed by hydrolysis.

copolymers (Kane *et al.* 1996). For instance, a two-phase polymer can be soaked in one reagent, which selectively absorbs into one phase. A cadmium salt could be taken up by a polyether phase. Subsequent treatment with hydrogen sulfide results in precipitation of cadmium sulfide within the polyether. While the volume fraction of sulfide formed is quite small, repeated cycles can give rise to higher fractions of the mineral.

Work on precipitation in lyotropic liquid crystalline amphiphile solutions has led to composite structures (see Sect. 4).

One characteristic of biological minerals is their elaborate shapes. For mechanical reinforcement of soft matrices, a filler should consist of high-aspect-ratio rod or plate particles. These may grow as a natural outcome of differing crystal growth rates along different crystal axes. However, most simple minerals are not sufficiently anisotropic to form such elongated particles. The shape of biological crystals is probably controlled by selective inhibition of growth on specific crystal faces. Very elongated silica particles are also formed in lipid vesicles, but it is not known how important these are for controlling shape in crystals. While synthetic methods offer many ways of controlling particle size, good methods for controlling shape are lacking. One route is to form long micromolds by phase separation in a two-phase polymer, which is cold-drawn to elongate the included phase. The included phase is then swollen with a metal alkoxide, which is hydrolyzed to oxide, Fig. 3 (Calvert 1996).

2. Toughness and Layered Structures

Large organisms operate in an environment that subjects them to fluctuating forces—from the action of wind and water on plants, and from the locomotion of animals. These fluctuating forces and collisions will often result in local damage that should not lead to catastrophic failure. As a result, many biological tissues contain structural features that add toughness without severely compromising stiffness or strength.

In mollusk shells and teeth, toughness may arise from added polymeric layers or from very fibrous structures (see *Shell: Properties*# and *Marine Teeth (and Mammal Teeth)*#). The addition of polymer could be especially effective if the polymer structure is capable of a large extension to break after yield, such as occurs in folded β-sheet proteins, where the unfolding of the chains leads to a large energy absorption, as has been discussed by Smith *et al.* (1999).

However, polymer layers will also lead to a significant loss of elastic modulus when compared to a wholly inorganic material.

Much effort has gone into increasing the serviceability of ceramic materials by improving their toughness. Much work was initially devoted to fibrous materials (see *Fibrous Ceramics*$^{\#}$), particularly silicon nitride, but the need for reliable, cheap processing methods was not met. Attention then switched to ceramic composites with fibrous reinforcements that would add toughness. These mostly proved to be unstable or reactive at the temperatures needed for turbine engines, while still being hard to process. Since then there have been many studies of ceramic/metal-layered structures, where the metal layers take the role of the polymers in mollusk shell. Such materials are of interest for use in engines and for armor. There is also interest in incorporating structured porosity into materials to improve fracture properties, or to modify elastic properties such as Poisson's ratio or piezoelectric response (Sigmund *et al.* 1998).

In the case of metal/ceramic-layered materials, the key problem is to absorb as much energy as possible within the metal layers. Hwu and Derby (1999) showed that the major energy absorption is due to metal drawing across the gap after the crack has passed. However, the extent of deformation occurring during the drawing was limited by the thinness of the metal layers. As a result, substantial increases in toughness were only seen when the volume fraction of metal was quite large. In contrast, mollusk shells show significant toughness with only a small percentage of polymer. It is likely that some combination of embedded metal, porosity, and interfacial debonding can give much enhanced toughness or impact resistance to ceramics. However, current methods for modeling fracture do not provide enough guidance for these designs and the biological models are only just being interpreted.

Recent studies on the influence of polymers on the crystallization of calcium carbonates *in vitro* has shown that metastable liquid complexes with anionic polymers can be formed which subsequently transform to calcite. Studies of carbonate biomineralization have shown that amorphous calcium carbonate does form either transiently or as a stable phase. It has also been shown that absorbed proteins do modify the fracture properties of carbonates. In particular, fracture surfaces show smooth conchoidal fracture, like glass, rather than the faceted fracture characteristic of normal crystals. This suggests that we have much to learn about the modification of the properties of crystals and amorphous solids by entrained polymer. In the example shown in Fig. 4, the polymer coating may serve to protect the surface from weakening cracks, and the entrained polymer may increase toughness. There are parallels with the Lanxide process for toughening alumina ceramic by entrained aluminum metal particles.

3. Polymers

The core difference between proteins and synthetic polymers is that protein synthesis provides total control of the sequence of units along a chain while the best-controlled polymerizations can only provide several blocks of functional units on a chain (see *Biosynthesized Materials: Properties and Processing*$^{\#}$). One obvious goal is the formation of synthetic polymers with enzyme-like catalytic activity. Many enzyme active sites have an array of active groups held in close proximity so as to interact with the substrate (target molecule), and to reduce the activation energy for reaction by a precise spatial array of ionic or hydrogen-bonding interactions. To achieve such a precise spacing of active groups on a synthetic polymer would require a rigid structure, which would in turn normally render the material insoluble and so inactive. Studies of dendrimer molecules (see *Dendrimers: Polymerization and Properties*$^{\#}$) with highly branched structures may lead to the required combination of flexible and soluble outer structures combined with a highly structured core. Many proteins also go through large changes in shape in response to the binding of substrates or other energy inputs. This seems to require a structure where two well-defined conformations are closely balanced in energy, so that the molecule can flip from one to another. Such a change is not likely in a wholly flexible system, and again some subtle combination of flexible segments and rigid units is required. There is still much to learn about these aspects of macromolecular design, but both the synthetic tools and the understanding are being acquired.

As discussed in the articles on silk (see *Silk Produced by Engineered Bacteria*$^{\#}$ and *Natural Protein Fibers**), structural proteins include regular and random regions in the chain structure that would be expected to give rise to crystalline and amorphous material (Calvert 1998). In contrast, the amorphous component of synthetic polymers arises from entanglements of the coiled chains that cannot be resolved during the crystallization process. New polymer properties may be achievable once polymer chains can be designed with such controlled sequences.

There has been a report of enzyme-like activity in a block co-polypeptide which enhances the rate of hydrolysis of tetraethoxysilane (TEOS; a standard reagent in sol-gel chemistry) as a suspension in water (Cha *et al.* 2000). If this block structure is made up of a sequence of units of a hydrophilic amino acid followed by units of a hydrophobic amino acid, it would be expected to be active at a water-solvent interface. The morphology of the silica that forms is dependent on the structure of the copolymer. This system is biomimetic both in the sense of employing a polypeptide catalyst and in the sense of it functioning in a multiphase system, since biological processes rarely occur in homogeneous solutions.

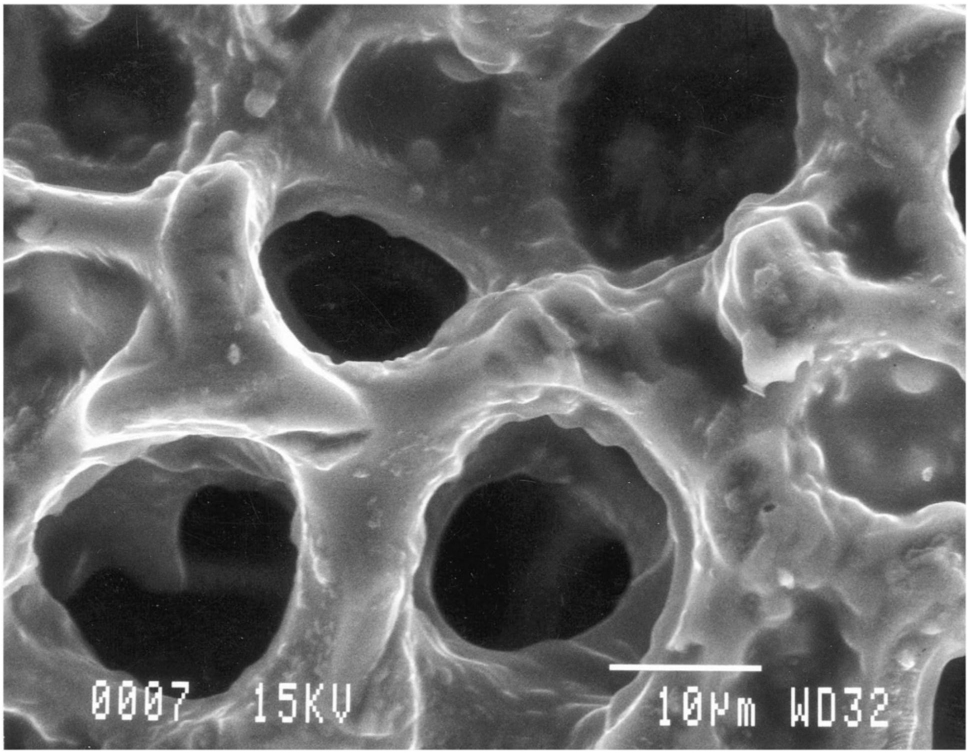


Figure 4
Porous interior of a sea urchin spine (calcite single crystal).

Many tissues, such as cartilage, arterial wall, and the walls of soft marine organisms, are swollen polymer structures. Swollen polymers, such as plasticized polyvinylchloride (PVC), do occur in artificial structures but they are usually avoided because loss of plasticizer leads to shrinkage and cracking. Even the swelling of wood with changes in humidity is a major impediment to its use in structures—though years of experience have taught us how to design around this problem. Skin does change in volume and properties as it takes up or loses water. The structure of amphibian skin keratin is apparently different from mammalian keratin for this reason.

There have been many suggestions that designers should make more use of soft structures. These could take the form of composites of hard fibers with rubbers. In which case, tires (National Materials Advisory Board 1994) and reinforced plastic tubing could be considered as examples. Liquid-swollen soft structures could also be used more in the synthetic world. For purely mechanical systems this may not make much sense, but in active systems such as batteries or muscle-like actuators, a liquid component is necessary and should probably be viewed as a soft material rather than simply as a liquid to be contained.

4. *Surfactants and Self-assembly*

Self-assembly is a hallmark of biological systems, including assembly of protein subunits into holoenzymes, of proteins and nucleic acids into virus particles, and of tropocollagen into collagen fibers. There has been increasing interest in synthetic self-assembly. In addition to the assembly of molecules with complementary hydrogen bonding to form supramolecular clusters (Lehn 1990), there are many papers on the assembly of charged polymers and particles to form multilayers (Bertrand *et al.* 2000), and on the assembly of particles or particles and films coated with complementary biological recognition molecules, such as the biotin-streptavidin system. Synthetic peptide self-assembly is discussed further in *Molecular Self-assembly*# and *Self-assembly by Phase Separation in Polymer Thin Films*#.

Work at the Mobil Corporation showed that mesoporous silica could be formed by hydrolysis of TEOS entrained in a highly concentrated water/silane/surfactant system. In this regime, the three-component mixture forms ordered structures with a range of symmetries. In one hexagonal phase, rods of water are surrounded by surfactant and embedded in a hydrophobic silane matrix. Hydrolysis of the silane under suitable conditions followed by drying and sintering, results in a porous silica with aligned pores of a few nanometers in diameter.

Growth of the lyotropic liquid crystal precursor is very sensitive to the environment. Ozin and co-workers have shown that complex particle morphologies can result from growth of these mesoporous structures in quiescent solutions as diffusion fields and surface forces interact (Yang *et al.* 1999). Several workers have shown how the direction of the rods or plates of silica can be controlled. Polymers can be introduced to form composite structures very reminiscent of some biological composites (Sellinger *et al.* 1998). This does seem to parallel the proposed importance of liquid crystals in the growth of many biological structures.

While one would expect that this approach could be extended to many other material combinations, the rules are not understood. Efforts to form similar structures, other oxides such as titania or various crystalline materials, have been only partly successful. Any rapid or localized conversion process possibly also disrupts the liquid crystalline organization.

Stupp and co-workers have produced a range of amphiphiles, which assemble into various ribbon and wedge structures, and have explored their catalytic activity (Stupp *et al.* 2000).

5. *Muscles*

There is a clear need for actuators for robot arms with properties comparable to human muscle. Unlike human muscle, an artificial muscle ought to respond to an electrical impulse by a change in length, since any chemical energy source must involve a system of pipes. There are many candidate materials, but all have considerable drawbacks at the moment. Piezoelectric actuators have a rapid response but the total strain is small, less than 0.1% and, as a result, complicated linkages are needed to obtain significant motion. Electrically conducting polymers show volume changes of a few percent, but this is still much below that needed for a muscle. Shape memory alloys are promising but the response is slow and the cycle life is limited. Motorized or hydraulic arms tend to be heavy and clumsy. A soft polymer actuator is the obvious analog for a muscle for fast but gentle manipulations.

Ionizable gels that respond to electrical fields and demonstration devices are discussed in *Synthetic Muscle*#. There is also much interest in composites of metal and ionic polymer as synthetic muscles (Shahinpoor *et al.* 1998).

6. *Processing Methods*

Freeform fabrication methods allow objects to be built as a series of layers directly from a three-dimensional computer representation. Current methods are described in *Freeform Fabrication*#. These allow only one material, or a material plus a soluble support structure, to be used at a time. It is clearly feasible to combine several different materials into a single object, which would allow the building of something much closer to an organism. A recent report on simple robots, which were allowed to evolve in a virtual environment and then were built by freeform fabrication, shows how evolutionary methods might be applied to manufacturing (Lipson and Pollack 2000).

To build a crude organism would require resolution at the scale of about 10 microns. Most current freeforming methods allow resolution down to about 100 microns and a limited materials set. Microcontact printing and related methods allow much higher resolution—down to about 1 micron but effectively are restricted to one layer (Xia and Whitesides 1998). There is much current interest in ink-jet printing methods which could provide the required 10 micron resolution while allowing many layers to be deposited (Fan *et al.* 2000).

In both synthetic and biological structures, it is useful to keep in mind the distinction between design and patterning. Phase separation, crystallization, and aggregation processes can give rise to patterns in two and three dimensions on a scale, from millimeters to nanometers, which reflects the kinetics of the separation and diffusion processes. To form working devices or organisms, we need to build to a nonrecurring design, which may include patterned elements. If we seek to adopt biomimetic processes, we will need to exploit self-assembling structures and patterns but within an overall design.

In current silicon technology, photolithographic methods can form two-dimensional designs down to less than 1 micron. Much finer resolutions are achievable in laboratory methods. Three-dimensional designs can be formed to below 1 micron using two-photon methods (Cumpston *et al.* 1999), but commercial freeforming methods are limited to about 100 microns resolution. In biology there are many examples of patterns forming at the nanometer level, but most designs are on the scale of individual cells, i.e., a few tens of microns. There are structures, such as sensing hairs, which are much finer but the spacing between them is still on the 10 micron scale.

7. *Cell Adhesion and Tissue Engineering*

Biomedical engineering is becoming more concerned with the problems of the long-term biocompatibility of synthetic implants. Material wear and degradation, and tissue loss or chronic inflammation due to changing mechanical loads will limit the lifetime of most implants to about 10 years. In addition, biocompatibility is not a material property but is very dependent on the specific environment of the implant—in terms of implant site in the body, animal species, and animal age. Strategies to eliminate these problems include the use of biodegradable materials that will eventually be replaced by natural tissue, and tissue-engineered implants where as much of the device as possible is formed from tissue grown *in vitro* on a synthetic support before implantation.

Tissue engineering imposes a need to understand the interaction between neighboring cells, and between cells and synthetic surfaces. Cell binding to a biological surface proceeds through a series of stages including physical adsorption, interaction between surface macromolecules and specific binding sites on the surface, reorganization of the macromolecules in the cell membrane to bring more binding sites into contact with the surface, and then specific changes in the cell induced by the surface.

At the physical adsorption level, studies of cell attachment to self-assembled monolayers on silicon or glass have shown that attached polyethylene oxide chains form a structureless hydrophilic barrier layer and will thus prevent adsorption. Different polar end groups allow cells to absorb, and printed surface patterns with binding and non-binding areas can be used to control cell shape (Zhang *et al.* 1999). Cells recognize and bind to simple short sequences of amino acids in a protein exposed on a surface. Arginine-glycine-aspartic acid (RGD) and arginine-alanine-aspartic acid-serine (RADS) are sequences that can be used to induce strong cell attachment (see *Molecular Self-assembly*#). Thus suitable polymers can be produced which will promote attachment of particular cell types on the surfaces for tissue engineering.

In the case of bone implants, there has been much interest in the use of bone morphogenic proteins (now produced from cloned bacteria). These constitute one group of a family of cell-signaling proteins known to induce the development of bone or other tissues when released locally in the body. A similar set of signals induces the growth of small blood vessels and is an important factor in the development of many cancers. By allowing implants to release such signaling proteins, it should be possible to speed up the integration of the implant. However, much remains to be learnt about the appropriate concentrations, gradients, and the timing of these signals. All these efforts take us in the direction of making synthetic organs that look more like a natural organ transplanted from an identical twin.

Belcher and co-workers (Whaley *et al.* 2000) have demonstrated that a phage library, displaying 10^9 different peptide sequences at the surface, can be used to identify short peptide sequences that selectively bind to inorganic semiconductor surfaces. Sarikaya and co-workers (Brown et al. 2000) have developed a similar method for the control of gold precipitation using *Escherichia coli* genetics.

8. Applications

Vogel has discussed the importance of biomimesis in engineering, asking whether it is possible to unequivocally attribute any engineering advance to a biological inspiration (Vogel 1998). There are no unambiguous examples and there are cases, such as attempts to imitate flapping flight, where a biological analogy may have inhibited progress. At the same time, many engineering advances clearly drew some inspiration from biology. In chemistry and materials, studies of the structure and properties of biological materials do suggest alternative approaches to particular problems and illustrate new properties that should be achievable in synthetic materials.

Biomimesis clearly will have an important role in new biomedical devices, and in new devices that try to combine biological structures or organisms with electronics in sensors and actuators. In structural materials, the obvious place for advance is the introduction of greater toughness into synthetic composites and ceramics. In electronics, there are not yet many signs of a conjunction between the hard, high-resolution, two-dimensional world of silicon and the soft, larger-scale, three-dimensional design of the brain and nervous system.

Bibliography

Bertrand P, Jonas A, Laschewsky A, Legras R 2000 Ultrathin polymer coatings by complexation of polyelectrolytes at interfaces: suitable materials, structure and properties. *Macromol. Rapid Commun.* **21**, 319–48

Brown S, Sarikaya M, Johnson E 2000 A genetic analysis of crystal growth. *J. Mol. Biol.* **299**, 725–35

Calvert P D 1994 Polymers for new materials. *Polymer* **35**, 4484–8

Calvert P D 1996 Biomimetic processing. In: Brook R J (ed.) *Materials Science and Technology*. VCH, Weinheim, Germany, Vol. 17B, pp. 51–82

Calvert P D 1998 Silk and sequence. *Nature* **393**, 309–10

Cha J, Stucky G, Morse D, Deming T 2000 Biomimetic synthesis of ordered silica structures mediated by block copolypeptides. *Nature* **403**, 289–92

Cumpston B, Ananthavel S, Barlow S, Dyer D, Ehrlich J, Erskine L, Heikal A, Kuebler S, Lee I, McCord-Maughon D, Qin J, Rockel H, Rumi M, Wu X, Marder S, Perry J 1999 Two-photon polymerization initiators for three-dimensional optical data storage and microfabrication. *Nature* **398**, 51–4

Fan H, Lu Y, Stump A, Reed S, Baer T, Schunk R, Perez-Luna V, Lopez G, Brinker C 2000 Rapid prototyping of patterned functional nanostructures. *Nature* **405**, 56–60

Hwu K L, Derby B 1999 Fracture of metal/ceramic laminates: I Transition from single to multiple cracking. *Acta Mater.* **47**, 529–43

Kane R S, Cohen R E, Silbey R 1996 Synthesis of PbS nanoclusters within block copolymer nanoreactors. *Chem. Mater.* **8**, 1919–24

Lehn J M 1990 Perspectives in supramolecular chemistry—from molecular recognition towards molecular information-processing and self-organization. *Angew. Chem. Int. Ed. Engl.* **29**, 1304–19

Lipson H, Pollack J B 2000 Automatic design and manufacture of robotic lifeforms. *Nature* **406**, 974–8

Mann S, Burkett S, Davis S, Fowler C, Mendelson N, Sims S, Walsh D, Whilton N 1997 Sol-gel synthesis of organized matter. *Chem. Mater.* **9**, 2300–10

National Materials Advisory Board 1994 Hierarchical Structures in Biology as a Guide for New Materials Technology. National Academy Press, Washington, DC

Sellinger A, Weiss P, Nguyen A, Lu Y, Assink R, Gong W, Brinker C 1998 Continuous self-assembly of organic-inorganic nanocomposite coatings that mimic nacre. *Nature* **394**, 256–60

Shahinpoor M, Bar-Cohen Y, Simpson J, Smith J 1998 Ionic polymer-metal composites (IPMCs) as biomimetic sensors, actuators and artificial muscles—a review. *Smart Mater. Struct.* **7**, R15–30

Sigmund O, Torquato S, Aksay I A 1998 On the design of 1-3 piezocomposites using topology optimization. *J. Mater. Res.* **13**, 1038–48

Smith B, Schaffer T, Viani M, Thompson J, Frederick N, Kindt J, Belcher A, Stucky G, Morse D, Hansma P 1999 Molecular mechanistic origin of the toughness of natural adhesives, fibres and composites. *Nature* **399**, 761–3

Stupp S, Pralle M, Tew G, Li L, Sayar M, Zubarev E 2000 Self-assembly of organic nano-objects into functional materials. *Mater. Res. Soc. Bull.* **25** (4), 42–8

Vogel S 1998 *Cats' Paws and Catapults: Mechanical Worlds of Nature and People*. Norton, New York

Whaley S R, English D S, Hu E L, Barbara P F, Belcher A M 2000 Selection of peptides with semiconductor binding specificity for directed nanocrystal assembly. *Nature* **405**, 665–8

Xia Y N, Whitesides G M 1998 Soft lithography. *Annu. Rev. Mater. Sci.* **28**, 153–84

Yang S, Yang H, Coombs N, Sokolov I, Kresge C, Ozin G 1999 Morphokinetics: Growth of mesoporous silica curved shapes. *Adv. Mater.* **11**, 52–5

Young J, Didymus J, Bown P, Prins B, Mann S 1992 Crystal assembly and phylogenetic evolution in heterococcoliths. *Nature* **356**, 516–8

Zhang S, Yan L, Altman M, Lassle M, Nugent H, Frankel F, Lauffenburger D, Whitesides G, Rich A 1999 Biological surface engineering: a simple system for cell pattern formation. *Biomaterials* **20**, 1213–20

P. Calvert
University of Arizona, Tucson, Arizona, USA

Bitumens: Modified

Bitumens are the residue products from the distillation of crude oil. They are used primarily for their waterproofing properties in the construction industry and their binding properties in the road industry. Bitumens are thermoplastic materials and have important limitations due to their temperature sensitivity. With the increasing demands of modern society, it is necessary to improve these properties where possible. In order to improve the properties relevant to applications, various forms of modification have been studied and introduced. These fall into the following categories: chemical modification, blending with other binder materials, and the incorporation of polymeric materials. Because of climatic differences throughout the world, the properties required for these modified bitumens can vary from country to country and region to region. Specifications are being drawn up by the major specifying bodies throughout the world for the polymer modified bitumens (PMBs), which parallel the corresponding specifications for the unmodified bitumens.

The nomenclature used in this article is as follows:

> *Bitumen* is the residue from the distillation of crude oil (US nomenclature: *asphalt* or *asphaltic cement*).
> *Asphalt* is a mixture of bitumen with graded aggregates (US nomenclature: *asphaltic concrete*).
> *Macadam* is a mixture of bitumen with single-size aggregates.
> *Lake Trinidad asphalt* is a naturally occurring product consisting of bituminous material with sand/fine aggregate.
> *Pitch* is the residue from the pyrolysis of coal tar.
> *Mastic asphalt* is a mixture of bitumen and limestone flour.
> *Cut back bitumens* are bitumens to which solvent oils have been added to reduce the viscosity (US nomenclature: *liquid bitumens*).
> *Coal tars* are the viscous liquid distillates from the distillation of coal. Their properties and composition depend on the distillation procedure and it is usual to signify the origin (e.g., coke oven tar, low-temperature carbonization tar).

This article also covers processes used to improve the handling properties of bitumens (e.g., emulsification in water) and the use of mix modifiers in bituminous materials (e.g., fibers).

1. Chemical Modification

The oxidation of bitumen by contact with air at an elevated temperature results in an increase in the softening temperature and a decrease in the penetration in an indentation test. The resultant properties depend on the grade of bitumen used as starting material and the actual conditions used. The products are termed *blown* or *oxidized* bitumens. The term *blown bitumen* is probably the best one, since the amount of oxygen added to the bitumen is very small. The main processes involved are dehydrogenation and polymerization (Corbett and Swarbrick 1960). Various catalysts have been proposed and ferric chloride, zinc chloride, and phosphorus pentoxide have been widely used.

The performance of a bitumen for some applications is related to its actual constitution. Bitumens are colloidal systems with highly aromatic material, asphaltenes, dispersed in a lower molecular weight medium, maltenes. The ease with which a bitumen can be emulsified in water is dependent on the degree of dispersion of the asphaltenes in the maltenes. Chemical additives (asphalt peptizing agents) are available that help to break down agglomerates of asphaltenes and disperse them in the maltenes. These include naphthenic acids and tall oil based acids. This type of treatment is particularly useful for improving the quality of emulsions that can be prepared from certain bitumens.

Another type of chemical addition treatment that gives improved strength of bituminous paving mixtures has been patented by Chemcrete Corporation (Latta and Leonard 1980). This process involves the addition of oil-soluble heavy metal salts (copper. manganese, cobalt) to the bitumen. These react with the bitumen and change its chemistry and molecular structure under the influence of temperature and oxygen.

The catalytic reaction results in the formation of ketones at the most reactive sites in the molecule, thereby reducing its susceptibility to oxidative aging and improving its antistripping properties. In a second consecutive stage, the organometallic components of the additive react with the ketones producing strong irreversible bonds between the bitumen molecules. This results in a bitumen with reduced temperature sensitivity The process is only suitable for use with "penetration grade" bitumens in coated stone preparation. The reaction does not start until the bitumen is spread in a thin film over a large surface accessible to oxygen. The process is relatively slow and the reaction is reported to be 80–90% completed in the course of the first 90 days after laying and compaction. However, there is some concern as to the extent of curing that can take place, since this could lead eventually to embrittlement of the binder.

2. *Blending with Other Binders*

2.1 *Coal Tar*

By careful selection of the origin of a tar component of low aromaticity, it is possible to prepare bitumen–tar blends that exhibit the best properties of the two components. Relative to bitumen, coal tar has better weathering properties and it shows better wetting of, and adhesion to, the surface of an aggregate. Coal tar also gives improved resistance to oil and petrol spillages. Although the coal tar industry is in decline, special areas for the application of bitumen–tar blends remain; for example, airport aprons and refuelling areas. Similarly, there are specific application areas for bitumen–pitch and bitumen–Lake Trinidad asphalt mixtures (BSI 1983, Dussek 1993).

2.2 *Sulfur*

The use of sulfur in bitumen–sulfur blends has been studied extensively, particularly as a means of improving resistance to permanent deformation of rolled asphalt wearing courses. The sulfur improves workability and increases dynamic stiffness as well as increasing resistance to permanent deformation. However, special equipment is needed to manufacture sulfur asphalt commercially and the mixing temperature requires very careful control (Slater and Zoumakadis 1982). Trials have been carried out using a combination of sulphur and rubber extended bitumen (Taman *et al.* 1989).

3. *Incorporation of Polymer Materials*

Extensive work has been carried out on the incorporation of polymeric materials into bitumen to modify the properties (Nicholls 1993, Read 1998, Bitume 1991). The following criteria have been put forward for road application (Denning and Carswell 1981). The polymer must

(i) be readily available;

(ii) resist degradation at asphalt mixing temperatures;

(iii) mix easily with bitumen;

(iv) improve resistance to flow at high road temperatures without making the binder too viscous at mixing and laying temperature or too stiff or brittle at low road temperatures; and

(v) be cost effective.

Most types of polymer have been examined and many are being used commercially. A considerable amount of information about the polymers being used, and the properties and performance of the resultant polymer-modified bitumens is available from the in-house literature issued by the leading polymer and bitumen producers and blenders. Specific information is available from the papers presented to the international conferences of organizations such as the European Bitumen Association (Eurobitume, Brussels, Belgium) and the Asphalt Institute (Lexington, KY, USA), and in the specialist road industry journals. (e.g., *Bitumen-Arbeitsgemeinschaft der Bitumen Industrie ev*, Hamburg, Germany: *Bitume Actualite*, Groupment Professionnel des Bitumes, Paris, France; and *Bulletin des Laboratoires des Ponts et Chaussees*, Paris, France).

3.1 *Thermosetting*

Two-component epoxy resins blended with bitumen display the properties of modified thermosetting resins rather than those of bitumen. The amount of bitumen that can be added to an epoxy resin system is limited since phase inversion can occur. This results in bitumen becoming the continuous phase with particles of cured resin dispersed in it, with the loss of the enhancing properties of the binder.

Since these thermosetting binders cure and harden by chemical action, they are not softened at high ambient temperatures. They are capable of resisting very high stresses imposed by traffic at difficult sites. The extended life of the binder justifies the use of a calcined bauxite as a surface dressing aggregate. This is a durable aggregate with an extremely high polished stone value (PSV). (PSV is a measure of the resistance of an aggregate to become smooth under the influence of traffic and of its retention of antiskid properties.)

Various proprietary epoxy bitumen–calcined bauxite systems have been developed and all display characteristic properties of exceptional strength, flexibility, and resistance to oil droppings and give essentially unchanged performance over a wide range of ambient temperatures.

3.2 Rubbers

Natural rubber has been used to modify bitumen for many years (Natural Rubber Producers Research Association 1964). Attention has now turned to synthetic rubbers (e.g., polybutadiene, polyisoprene, butyl rubber, neoprene, and ethylene propylene diene monomer (EPDM)). Recovered tyre crumb is also used.

(*a*) *Natural rubber*. This is used in both rubber crumb and latex form, the latex being used particularly in the preparation of rubberized bitumen emulsion and in asphalt/macadam by direct addition to the mixing box. The use of rubber crumb is limited since it tends to degrade at the temperatures needed to incorporate it into bitumen. Although the available natural latex is normally in anionic form, rubber latex in a form that is compatible with cationic systems is now available commercially or can be made *in situ* by the use of charge reversal agents.

(*b*) *Recovered tyre crumb*. This has been used in a number of proprietary processes, particularly in France and Belgium. In one process, the finely pulverized tyre crumb (18–20%) is blended with bitumen (77–79%) and an aromatic flux oil (3%) at 205 °C. Optimum performance is obtained when a mixture of equal parts of natural and synthetic rubber is used. Careful selection of the rubber crumb is needed in order to prevent the formation of a heterogeneous binder in which the rubber acts as a flexible filler (see Elastomers, Recycling of[#])

(*c*) *Synthetic rubbers*. These are used mainly in the form of latex for addition to bitumen emulsions. They are also used to modify bitumen for asphalt production (Heather *et al*. 1988). The latex is introduced as a separate feed into the mixing box of the asphalt plant.

(*d*) *Thermoplastics*. The incorporation of thermoplastics (e.g., polyethylene, polypropylene, and EVA copolymers from the copolymerization of ethylene with vinyl acetate) into bitumen tends to increase the stiffness and viscosity of the bitumen at normal service temperatures. They do not, however, confer improved resilience (Denning 1984).

(*e*) *Thermoplastic rubbers*. Thermoplastic rubbers of the type typified by the SBS (styrene–butadiene–styrene) block copolymers of styrene and butadiene behave as a thermoplastic during hot processing and as a vulcanized rubber at ambient temperatures (Vonk 1989). Both linear and radial copolymers are used. The polybutadiene and polystyrene portions are incompatible and form a two-phase system consisting of domains of glassy polystyrene interlinked by flexible polybutadiene chains. The resulting network structure behaves as if the molecules were chemically cross-linked or vulcanized and exhibits high tensile strength (Read 1998).

The network structure can be disrupted by addition of solvent or by heating the polymer to temperatures above the melting point of the polystyrene. The network, with all of its original properties, can be restored by eliminating the solvent or by cooling the polymer to below the glass transition point of polystyrene. The tensile strength of the butadiene–styrene block copolymer depends on the styrene content, molecular weight, and temperature.

The suitability of a bitumen for modification in this way is related to its composition, in particular the percentage of asphaltenes, and to the percentage of aromatics in the maltenes fraction of the bitumen. During the dissolution of the rubber, the following phenomena occur: polymer crumbs, molten at high temperature, are sheared into small droplets and the molecules start to uncoil and disentangle, assisted by the oils and resins in the bitumen.

Depending on the concentration of rubber, a continuous bitumen or a continuous rubber phase will develop. Inversion occurs at a rubber content of 6–8%. Mixtures for road application are of the former type (continuous bitumen phase). Mixtures for roofing purposes belong to the second type (rubber network).

Good mixes exhibit the following properties:

(i) high increase in softening point;

(ii) lowering of the Fraas point (this is the temperature at which a bitumen first becomes brittle as indicated by the appearance of cracks when a thin film of the bitumen on a metal plaque is cooled and flexed in accordance with specific conditions);

(iii) high values of tensile strength and elongation: and

(iv) high cohesivity when used with aromatic oils for the improvement of low-temperature behavior.

In addition to the SBS copolymers, other systems used include SB (styrene–butadiene), SIS (styrene–isoprene–styrene), and SIBS (styrene–isoprene–butadiene–styrene), either alone or in admixture with SBS.

An alternative to the use of block copolymers described above is the incorporation into the bitumen of a low molecular weight rubber, followed by vulcanization *in situ* (e.g., ELF Union 1979, McKenna 1982). Bonding takes place between the polymers and the bitumens to produce homogeneous products with the desired increased stiffness at high temperatures and flexibility at low temperatures, combined with good cohesive strength and adhesive power. During the process, active sites in the bitumen are taken up. This results in improved resistance to oxidative aging and an increase in the working life.

3.3 Incorporation of Polymers into Bitumen

The conditions required for the incorporation of polymers into bitumen depend on the particular polymer, but normally high temperature and high shear are needed. In certain cases, diluent oils can be used to facilitate the blending, but this is not possible in every case owing to compatibility limitations. It is possible to use a polymer in latex form, although direct addition of a latex to hot bitumen must be carefully controlled to avoid frothing. Latexes, however, are used extensively in the preparation of polymer-modified bitumen emulsions. The latex may be incorporated into a bitumen emulsion by mixing the two together provided that the emulsion and the latex are ionically compatible. Alternatively, the polymer latex may form part of the aqueous-phase soap solution used in the emulsification of the bitumen.

3.4 Other Additives

Although strictly they do not modify the properties of bitumens, there are a number of additives which are used to enhance performance of the binders or bituminous products made from them.

(*a*) *Adhesion agents* (*antistripping agents*). These are cationic chemicals with long chain or cyclic amine or amidoamine structures that assist the bitumens to adhere strongly to the surfaces of stones. Also they alter the surface tension of the bitumen, which increases the contact angle between the bitumen and the stone. This helps to resist the stripping of the bitumen from the stone surface by the action of water. The use of silanes as adhesion agents/antistripping agents has also been examined.

(*b*) *Fibers*. These tend to form a network structure within the bitumen with an increase in viscosity. Thus they allow for increased loading of bitumen on to the aggregate in hot-mix processes such as porous asphalt (whispering asphalt) and stone mastic asphalt (SMA). Fibers are also used in the slurry seal process to give thicker slurries. Cellulose fibers, recovered newspaper fibers, acrylic fibers, glass fibers, etc., are used in the road industry. Fibers are also used as replacement for asbestos in certain bituminous products for the construction industry.

(*c*) *Carbon black, silica*. These have been used as fillers in hot-mix processes to give improved wearing properties. This is analogous to their use in motor vehicle tires. It is believed that carbon black helps to reduce oxidative embrittlement of the surfaces.

(*d*) *Gilsonite*. This is a naturally occurring asphaltene found in Utah, USA. It is used to harden and strengthen bitumens for certain uses.

(*e*) *Waxes*. Waxes have been incorporated into bitumen for specific applications: for example, molten wax/bitumen mixtures to give a zero penetration binder that can be emulsified. The resultant emulsion is mixed with peat and pressure molded to form plant pots for the horticultural industry.

(*f*) *Amides*. The use of ethylene-bis-stearamide (EBS) to increase the softening point has been made. It is possible that the EBS may be beneficial as a slip agent in the bituminous backing of carpet tiles.

4. *Applications of Modified Bitumens*

The main applications of modified bitumens are in the road and roofing industries and the amounts are increasing annually.

4.1 Roofing

Oxidized or blown bitumens are used extensively in roofing applications, although the use of polymer-modified bitumens in roofing felt manufacture is increasing rapidly. The polymers used are mainly atatic polypropylene, EPDM, and SBS copolymers. The level of addition of the polymer is normally high (greater than 13%).

Emulsions of polymer-modified bitumens are also finding increasing application in the construction industry owing to the improved properties of the resultant film over those of unmodified bitumens. One such application is the spraying of an emulsion and a breaking agent through a twin-jet spray to produce a seamless membrane on a roof, for example.

4.2 Roads

The main growth area of bitumens in road applications is the use of polymer-modified bitumens in both road construction and road maintenance. The market for modified bitumens grew by 20% between 1987 and 1988 and continues to grow. Of the different polymer types used, styrene thermoplastic rubbers showed the greatest increase. A breakdown of the West European polymer modified road binders market for 1987 and 1988 is given in Table 1.

The application areas fall into three main categories: asphalt/macadam, surface dressing, and slurry seal. The availability of polymer-modified bitumens has made possible the new generation of thin and ultrathin surfacings which are being increasingly used to repair worn surfaces. These processes also require polymer-modified bitumens to be used in the tack or bond coat.

Polymer-modified bitumens are finding increasing application in SMA, in part by virtue of the fact that the increased binder viscosity allows for higher loadings of the binder needed on the aggregate than would be possible with the unmodified bitumens.

For this application polymer-modified bitumens compete with the use of fibers in the mix (Davies 1999, Milster 1993).

(*a*) *Asphalt/macadam*. In order to offset the deformation and cracking due to increasing traffic density and axle weights, modified bitumens are becoming more widely used in new road construction and road maintenance.

EVA–bitumen systems are used extensively to give improved resistance to flow at elevated temperatures. This is particularly useful when rutting is a problem. SBS–bitumen systems give greater flexibility and are particularly useful in preventing cracking. The use of polymer-modified bitumens enables thinner wearing courses to laid if desired.

Table 1
West European polymer-modified bitumen road market.

Polymer	Share of market (%) 1987	1988
Styrene–thermoplastic rubber	27	44
Styrene–butadiene diblock copolymer	17	13
Ethylene vinyl acetate (EVA)	25	11
Reclaim rubber	10	9
EPDM	21	12
Polyolefins		10
Others		1

Source: Shell International.

The amount of polymer used in asphalt mixes is in the region of 3–7% by weight of binder. Normally, the EVA or SBS is incorporated into the bitumen at elevated temperature to give the modified binder. It is possible, however, to use the polymer in latex form (synthetic rubber or natural rubber) and to add this at the actual stone coating stage (Heather *et al.* 1988). The water present in the latex flashes off during the mixing stage.

One area in which polymer-modified bitumens play an important part is in pervious macadams. These are open-textured wearing courses that allow water to penetrate and to run off sideways through the wearing course (Hoban *et al.* 1986, Coldwell and Daines 1989). This gives a considerable reduction in spray in wet weather.

Additionally, because of the open texture, there is a marked reduction in noise level. This has led to these surfaces being referred to as "whispering asphalt." The modified bitumens are needed in order to give greater strength and allow the higher binder contents needed. Various polymers have been used including SBS, EVA, and natural and synthetic rubbers. Fibers can also be used in pervious asphalt.

Polymer-modified bitumens are used in the tack (or bond) coats for pervious macadams in order to prevent the water entering the base course.

(*b*) *Surface dressing*. In the USA, surface dressing is called chip sealing. This is a process for renewing the wearing surface of a road by applying a thin coating of hot bitumen (cut back or emulsified) to the surface, followed by the application of aggregate chippings. The availability of modified bitumens has enabled the process of surface dressing to be extended to more difficult sites than can be coped with by unmodified bitumens (Hoban 1990). The modified bitumens can be applied in both cut back and emulsified form (Brule 1995). The preferred polymers are thermoplastic rubbers of the SB and SBS type and proprietary products produced by *in situ* vulcanization/chemical linking of low molecular weight rubbers dissolved in the bitumen.

In addition to the reduced temperature sensitivity and increased strength of the binder, the elastic properties of these binders confer good cohesive strength and adhesive power (i.e., they hold the aggregate chips more effectively). This gives good initial chip holding and resistance to applied strain. For very difficult sites, such as roundabouts, major road junctions, and approaches to traffic lights and other danger points, surface dressing based on the bitumen extended epoxy resin–calcined bauxite systems is the preferred treatment.

(*c*) *Slurry seal/microsurfacing*. The use of polymer-modified bitumens in the slurry seal process has greatly increased the scope of this type of road maintenance system (this is a process for sealing a road surface and/or renewing the wearing surface involving application of a slurry of fine aggregate and bitumen emulsion as a thin surface layer). Systems based on emulsions prepared from polymer-modified bitumens are used, as are those in which the polymer (natural or synthetic) is introduced in the form of latex. The latex can be incorporated into the aqueous phase at the preparation of the emulsion or may be blended into the bitumen emulsion prior to mixing with the aggregate.

With polymer-modified emulsions, larger-size aggregates may be used and the slurry placed in greater thickness (sometimes referred to as microsurfacing). This, combined with the better high-temperature stability and low-temperature flexibility, has enabled the scope of slurry seal to be extended to more heavily trafficked roads than is possible with slurries based on unmodified bitumen emulsion.

Slurry seal systems based on natural rubber, SBS, neoprene, EVA, etc., are operated under proprietary names, as are also systems which additionally contain fibers.

The suitability of a specific polymer depends on the polymer–bitumen–aggregate interactions. As a consequence, systems tend to be tailormade to suit the available materials. Attempts have been made to bring some order, and the International Slurry Seal Association published "Recommended Performance Guidelines for Polymer Modified Microsurfacing" AI43 in May 1987. In addition to giving recommended test methods and procedures, this publication gives a requirement of 2.5% polymer solids based on bitumen weight.

(*d*) *Miscellaneous*. Polymer-modified bitumens find application in a range of miscellaneous applications including mastic asphalt, joint sealants, waterproofing membranes, bridge deck membranes, stress absorbing membrane interlayers, etc.

5. Test Methods for Modified Binders

In order to determine the type and concentration of polymer present in a modified binder, use in generally made of IR spectroscopy. The morphologies of some bitumen–polymer blends can be determined by optical microscopy techniques with fluorescence. Under the test conditions, bitumen is not fluorescent, whereas plastic rubbers, for example, emit a yellow-green fluorescence. So far no standard performance tests have been adopted but the following are included in those under consideration:

(i) Mini fretting test: this determines the stone lost from a surface dressing produced and tested under standard conditions;

(ii) toughness and tenacity test: this determines the force needed to pull a steel ball from a sample under standard conditions;

(iii) ring-and-ball test: this measures the softening point;

(iv) penetration: this determines the hardness by measuring the penetration of a standard needle under standard conditions;

(v) Fraas brittleness point test;

(vi) elastic recovery;

(vii) hardening in the rolling film oven test;

(viii) storage stability; and

(ix) measurement of plasticity range.

See also: Roofing Materials*

Bibliography

Bitume 1991 Various papers. *Groupement Professionnel des Bitumes* **92**, 5–36

Brule B 1995 Les emulsions modifiees. *Route Actualite* **41**, 39–47

BSI 1983 *Bitumens for Building and Civil Engineering*: part 3. *Specifications for Bitumen Mixtures*, BS 3690. British Standards Institution, London

Coldwell D M, Daines M E 1989 Progress in the trials of pervious macadam. *Highways* **57** (1945), 15–8

Corbett L W, Swarbrick R E 1960 *Assoc. Asphalt Paving Technol.* **29**, 104

Davies A 1999 The use of stone mastic asphalt and high modulus base on the M1-A1 link road DBFO project. *The Asphalt Year Book 1999*. Institute of Asphalt Technology, London, pp. 83–6

Denning J H 1984 New binders. *J. Inst. Asphalt Technol.* **35**, 40–4

Denning J H, Carswell J 1981 Improvement in rolled asphalt surfacings by the addition of organic polymers, TRRL Report 989. Transport and Road Research Laboratory, Crowthorne, UK

Dussek I J 1993 Modifying bitumen with Trinidad Lake asphalt. *SCI Lecture Papers* (paper no. 0033). Society of Chemical Industry, London

ELF Union 1979 Process for preparing bitumen polymer compositions. Br. Pat. 1,548541

Heather W, Laitinen J T, Rowe G M 1988 The asphapol process and polyphalt materials. *Highways* **56** (1933), 48–52

Hoban T W S 1990 Modified binders for surface dressing. *Chemistry & Industry*, 3rd September, 538–43.

Hoban T W S, Liversedge F, Soarby R 1986 Recent developments in pervious macadam surfaces. *Quarry Manager* **13** (8), 35–6

Latta L, Leonard J B 1980 High strength modified asphalt paving compositions. US Pat. 4 234 346

McKenna M 1982 Taking a fresh look at our crumbling roads. *Surveyor* (28 Jan.), 16–17

Milster R 1993 Herstellen und Einbauen von Splittmastix Asphalt. Ein Erfahrungsbericht. *Bitumen* **55**, 146–8

Natural Rubber Producers Research Association 1964 Natural rubber in road surfacing. *Tech. Bull. No. 9*. NRPRA, London

Nicholls J C 1993 Generic types of binder modifier for bitumen. *SCI Lecture Paper* (paper no. 0035). Society of Chemical Industry, London

Read J 1998 On the road. *Chem. Br.* **34**, 46–9

Slater R J, Zoumakadis I 1982 Some properties of sulphur bitumen mixes. *Highw. Public Works* **50** (April), 13–4

Taman A R, Helmy M A, Mekkaw A 1989 Utilisation of sulphur and rubber in modifying bitumens for road materials. *J. Inst. Asphalt Technol.* **41**, 61–6

Vonk W 1989 Polymer modification of bitumens. *J. Inst. Asphalt Technol.* **41**, 21–6

P. Nayler
Nayler Chemicals Ltd., Wigan, UK

Block Copolymers as Precursors to Porous Materials

Incorporating voids into a polymer in a controlled manner offers unique opportunities for applications in foams, membranes, filters, and microelectronic materials. Each application places demands and restrictions on both the polymer matrix and the nature of the voids. The polymer must have a high enough modulus to support the void structure while maintaining other characteristics such as thermal stability or resistance to specific solvents. Issues arise as to whether the pores are open- or closed-celled, and the size and number density of pores. As the pore size decreases, the force to collapse the pore increases, which places severe demands on the matrix polymer, particularly in terms of the mobility of the polymer over small distances.

There are several strategies for the preparation of porous materials (Khemani 1997). One uses a gas directly as a blowing agent (Shutov 1991). In a later method, a polymer is swollen with supercritical fluid (SCF), the pressure is reduced, and the SCF converts to a gas, foaming the polymer (Beckman 1995). Alternatively, a gas can be produced via a chemical reaction, whereby a gaseous by-product is generated. A different route is based on emulsion, micellization, or dispersion techniques (Ruckenstein 1997). Here, the dispersed component consists of a volatile liquid or polymer with the supporting phase comprised of a monomer, which forms a rigid matrix on polymerization. A porous medium is obtained by removal of the dispersed phase via evaporation, extraction or degradation by thermal, chemical, or photolytic means. An advantage of this approach is that the system can self-assemble into particular structures such as body-centered cubic arrays, hexagonally close-packed cylinders, or bicontinuous structures. Another approach is to use a thermally, solvent-, or chemical-induced phase separation of a polymer with a second component that is either a volatile liquid or a material that can easily be degraded (Kiefer *et al.* 1999). The size and spatial arrangements of the voids will be dictated by the morphology generated in the phase separation process.

However, for many advanced applications, the typical size scales of the pores are too large. In microelectronics, where feature sizes are at one-quarter micron level, it is imperative that the size of the pores produced be much smaller than this. In addition, a closed-cell pore structure is essential and the porous material must have high thermal stability, high modulus, and the lowest possible density.

One route that satisfies these criteria is through the use of block copolymers. Block copolymers self-assemble into well-ordered arrays of nanoscopic domains with domain sizes defined by the length of the chain (see *Block Copolymer Phase Behavior*[#]). If one of the blocks can be degraded thermally or photolytically and the flow of second block can be arrested, a porous material can be produced where the size, shape, and spatial distribution of the pores is dictated by the original morphology of the copolymer. Whether the application be in the bulk or in thin films, key issues arise concerning the spatial organization and orientation of the pores. Here, several different approaches in the use of block copolymers for the generation of nanoporous media will be discussed.

1. Thermal Approach

One approach to nanoporous materials is to couple a thermally stable polymer, such as aromatic polyimides or poly(phenylquinoxaline)s, to a thermally labile polymer such as poly(methyl methacrylate) (Hedrick *et al.* 1993, 1999). The thermally labile polymers can be prepared via living anionic, ring opening, group transfer or controlled free radical polymerization methods (see *Block Copolymer Synthesis*[#]). In general, the thermally stable block is prepared by a condensation polymerization as, for example, in the case of polyimides, which are obtained from the cycloimidization of either a poly(amic acid) or poly(amic acid ester). By functionalization of each of the chains, the polymers are coupled to form a multiblock copolymer, which undergoes microphase separation.

Thin films are spincoated, the casting solvent is removed, and the film is annealed to promote the microphase separation of the blocks. The matrix polymer is then cycloimidized well below the decomposition temperature of the labile component. Complete conversion to the polyimide creates a matrix that has thermal stability and is partially ordered. Subsequently, the copolymer is heated to a temperature below the glass transition temperature of the matrix but above the decomposition temperature of the labile component. Degradation of the labile block occurs, volatilizing the components, leaving behind a porous polyimide.

There are several features to this approach that can lead to a collapse of the porous structure or to cell sizes that are too large. If degradation occurs by a depolymerization or unzipping, the rate of degradation, reflected in the zip length and ceiling temperature, must not be too high. For example, the depolymerization of poly(alpha-methylstyrene) to monomer occurs more rapidly than the diffusion of the monomer out of the film. The monomer volatilizes in the matrix, forming exceptionally large pores. Poly(methyl methacrylate), on the other hand, depolymerizes less rapidly and a nanoporous structure can be generated. If degradation occurs by chain scission, the size of the fragments must be sufficiently small to allow diffusion out of the film.

Diffusion of the degradation products through the matrix poses another challenge. For diffusion to occur, the degradation products must be soluble in the matrix and, then, must permeate through the matrix. This will lead to a plasticization of the matrix and a reduction of its T_g. If the T_g is depressed sufficiently, the matrix will flow leading to a collapse of the pores (Hedrick *et al.* 1993, 1999).

Finally, the volume fraction of the labile component must be sufficiently low so that the pores are closed-cell, i.e., not interconnected. Typically, this requires that the volume fraction of the labile component be less than 0.3. Void formation can be studied *in situ* using real-time, small angle x-ray scattering SAXS and infrared absorption. Both techniques have been used effectively to monitor the kinetics of pore growth in polyimide films (Sanchez *et al.* 1995, Cha *et al.* 1996, Fodor *et al.* 1997). The ultimate test of the success of void formation is the reduction in the dielectric constant. Polyimides have a typical dielectric constant of 2.9. Using the thermal approach described above, dielectric constants have been reduced to 2.2, representing an incorporation of 23% voids by volume into the polymer while retaining a closed-cell, nanoscopic void structure.

2. *Vapor and Irradiative Degradation*

An alternative approach to the generation of nanoporous media using block copolymers is to expose the copolymer to a corrosive vapor that will simultaneously degrade one of the blocks while cross-linking the other. Chaikin and co-workers (Park *et al.* 1997) demonstrated the effectiveness of this approach in a study on block copolymers of polystyrene and polyisoprene denoted P(S-b-I). Using an asymmetric P(S-b-I) copolymer containing spherical polyisoprene domains, exposure to ozone cross-linked the polystyrene matrix while degrading the polyisoprene block. This left a film of cross-linked polystyrene having spherical air cavities with a morphology that was identical to that of the parent film. Subsequent, reactive ion etching of this film transferred a relief of microdomain structure into the substrate (see Block Copolymer Templates for Nanolithography[#]).

Thomas and co-workers (Avgeropoulos *et al.* 1998) treated thin films of poly(pentamethyldisilylstyrene) P(PMDSS)/PI copolymers with ozone coupled with UV irradiation. The initial copolymer had the bicontinuous gyroid morphology. Upon ozone/UV exposure, the PI block was degraded and the P(PMDSS) phase was cross-linked and converted to a silicon oxide. Consequently, a porous glass is produced having a structure similar to that of the original copolymer with void volume fractions of up to 0.51. Hashimoto and co-workers (Hashimoto *et al.* 1997) had previously performed similar studies using P(S-b-I) having a gyroid structure. These co-workers subsequently coated the structure with nickel using electroless deposition methods generating a nanohybrid structure.

Asymmetric, diblock copolymers of polystyrene and poly(methyl methacrylate), PMMA were investigated by Russell and co-workers (Thurn-Albrecht *et al.* 2000) where the cylindrical PMMA microdomains were oriented normal to the surface thin films using an electric field. Upon exposure to deep ultraviolet radiation, the PMMA was degraded and the PS matrix was cross-linked yielding a hexagonally-packed array of nanopores perforating the film. These films are being used as templates in which to grow nanowires and as etching masks for pattern transfer to the underlying substrate.

3. *Partial Degradation*

Rather than decomposing an entire block of a copolymer, only a fraction can be removed. Ding and Liu (1998) used diblock copolymers of poly(*t*-butyl acrylate), PTBA, with poly(2-cinnamoylethyl methacrylate), PCEMA, having PtBA cylinders. Reaction of the copolymers with trimethylsilyl iodide replaced the *t*-butyl group with a trimethylsilyl group. Exposure of the bulk copolymer to UV cross-linked the PCEMA block, freezing in the copolymer structure. Subsequently, the trimethylsilyl acrylate block is hydrolyzed, generating polyacrylic acid (PAA). The loss of the trimethylsilyl group causes a reduction in the occupied volume of this block. Upon drying, the PAA block shrinks, coating the interior walls of the cylinders in the cross-linked PCEMA matrix. PAA is water soluble, and consequently water can flow through such a membrane. However, the flow can be adjusted by a change in the pH or by the addition of Ca^{2+} ions, thereby generating a nano valve to regulate flow.

4. *Mesoporous Silica*

The use of amphiphilic block copolymers to direct the organization of polymerizing silica has been shown to produce well-ordered hexagonal mesoporous silica with pore sizes of 75–300 Å (Zhao *et al.* 1998, Beck *et al.* 1992, Krämer *et al.* 1998). Unlike conventional surfactants, block copolymers offer the possibility of fine-tuning the polymer/solvent phase behavior by adjusting molecular weight, composition, and macromolecular architecture. Both ionic and nonionic amphiphilic block copolymers have been used to produce mesostructured silica with the ultimate objective of producing porous systems upon calcination. For example, nonionic alkyl poly(ethylene oxide) oligomeric surfactants and poly(alkylene oxide) block copolymers in acidic media have been used to produce highly ordered mesoporous silica structures (Zhao *et al.* 1998). The periodic assembly of the organic-inorganic hybrids occurs by a hydrogen-bonding pathway involving cationic silica species that are present in strong acidic conditions. Alternatively, block copolymers with one polyelectrolyte block can be used as templates in a sol-gel process. Cationic poly butadiene-poly(vinyl pyridinium) and anionic poly(ethyl ethylene)-polystyrene sulfonate block copolymers have been used (Krämer *et al.* 1998). The pore size and connectivity in the silica gel is an exact copy of the original self-assembled structures (pore sizes 10–50 nm). Furthermore, the curvature and related aggregation structure of the ionic block copolymer could be modified by adding low molecular weight salts.

5. *Summary*

The generation of nanoporous media from block copolymer requires an easy means of eliminating or partially eliminating one component. The surface energy of the matrix coupled with the large increase in the surface area produces a strong driving force to collapse the structure. This mandates that the matrix component be able to overcome this force, which can be accomplished by cross-linking or ordering the matrix component. By a judicious selection of the blocks or by manipulating the processing conditions, nanoporous structures can be prepared. The applications of such structures are numerous, though the impact of nanoporous media has only recently begun to be felt.

Bibliography

Avgeropoulos A, Chan U Z-H, Lee V Y, Ngo D, Miller R D, Hadjichristidis N, Thomas N L 1998 Synthesis and morphological behavior of silicon-containing triblock copolymers for nanostructure applications. *Chem. Mater.* **10**, 2109

Beck J S, Vartuli J C, Roth W J, Leonowicz M E, Kresge C T, Schmitt K D, Chu C T W, Olson D H, Sheppard E W, McCullen S B, Higgins J B, Schlenker J L 1992 A new family of mesoporous molecular sieves prepared with liquid crystal templates. *J. Am. Chem. Soc.* **114**, 10834

Beckman E J 1995 Polymer nanostructures via critical fluid processing. In: Vincenzini P (ed.) *Adv. Sci. Technol. New Horizons for Materials, Techna Srl*, 151

Cha H J, Hedrick J L, DiPietro R, Blume T, Beyers R, Yoon D Y 1996 Structures and dielectric properties of thin polyimide films with nanofoam morphology. *Appl. Phys. Lett.* **68**, 1930

Ding J, Liu G 1998 Semi-shaved and fully shaved hollow nanospheres from polyisoprene-*block*-poly(2-cinnamoylethyl methacrylate). *Chem. Mater.* **10**, 537

Fodor J S, Briber R M, Russell T P, Carter K R, Hedrick J L, Miller R D 1997 Transmission electron microscopy of 3F/PMDA-polypropylene oxide triblock copolymer-based nanofoams. *J. Poly. Sci. Poly. Phys. Ed.* **35**, 1067

Hashimoto T, Tsutsumi K, Finaki Y 1997 Nanoprocessing based on bicontinuous microdomains of block copolymers—nanochannels coated with metals. *Langmuir* **13**, 6869

Hedrick J L, Ladadic J, Russell T J, Hofer D, Wakharker V 1993 *Polymer* **34**, 4717–26

Khemani K C (ed.) 1997 *Polymeric Foams: Science and Technology*. American Chemical Society, Washington, DC

Hedrick J L, Carter K R, Labadie J W, Miller R D, Volksen W, Hawker C J, Yoon D Y, Russell T P, McGrath J E, Briber R M 1999 *Adv. Polym. Sci.* **141**, 1

Kiefer J, Hedrick J L, Hilborn J G 1999 Macroporous thermosets by chemically-induced phase separation. *Adv. Polym. Sci.* **147**, 161

Krämer E, Förster S, Göltner C, Antonietti M 1998 Synthesis of nanoporous silica with new pore morphologies by templating the assemblies of ionic block copolymers. *Langmuir* **14**, 2027

Park M, Harrison C, Chaikin P M, Register R A, Adamson D H 1997 Block copolymer lithography—periodic assays of $\sim 10^{11}$ holes in one square centimeter. *Science* **276**, 1401

Ruckenstein E 1997 Concentrated emulsion polymerization. *Adv. Polym. Sci.* **127**, 1

Sanchez M I, Hedrick J L, Russell T P 1995 *J. Polym. Sci. Part B Polym. Phys.* **33**, 253–7

Shutov F A 1991 Blowing agents for polymer foams. In: Klempner D C, Frisch K C (eds.) *Handbook of Polymeric Foams and Technology*. Hanser, München, Chap. 17

Thurn-Albrecht T, Steiner R, DeRouchey J, Stafford C M, Huang E, Bal M, Tuominem M, Hawker C J, Russell T P 2000 Nanoscopic templates from oriented block copolymer films. *Adv. Mater.* **12**, 787–91

Zhao D, Feng J, Huo Q, Melosh N, Fredrickson G H, Chmelka B F, Stucky G D 1998 Triblock copolymer synthesis of mesoporous silica with periodic 50–300 Å pores. *Science* **279**, 548

T. P. Russell[a] and J. L. Hedrick[b]
[a] *University of Massachusetts, Amherst, USA*
[b] *IBM Almaden Research Center, San Jose, California, USA*

Block Copolymers as Templates for Functional Materials

There are applications and devices which require controlled distribution of material functionality (electrical, optical, catalytic, magnetic) in two or three dimensions. At the nanometer length scale, attempts to meet this challenge have included template-mediated materials chemistry (Martin 1994) in which track-etched membranes, porous alumina, and zeolites serve as the nanoscale reaction vessels for the synthesis of the functional materials. The ability to control both the length scale and the spatial organization of block copolymer morphologies (see *Block Copolymer Phase Behavior*[#]) makes these materials particularly attractive candidates for use as templates in the synthesis of functional nanocomposites. Appropriate choices of the repeat units of the block sequences render them capable of selectively sequestering preformed inorganic nanoclusters or selectively solubilizing inorganic reagents for *in situ* cluster synthesis. Methods exist to produce nanoscale voids (see *Block Copolymers as Precursors to Porous Materials**) which percolate through the structure, leading to processes which coat or backfill the channels with functional materials.

1. Block Copolymer Nanoreactors

The *in situ* production of inorganic nanoclusters within nanoscale-dimensioned block copolymer domains can eliminate preliminary steps to synthesize, stabilize, and store the clusters, and cumbersome mixing steps required to make nanocomposite films with uniform dispersion. The capacity for producing long-range alignment (see *Block Copolymers, Flow Alignment of*[#]) means that patterned arrays of the clusters are readily achievable. In Fig. 1 a spherical domain of a block copolymer is represented to contain receptor groups, in this case carboxylic acids capable of binding positive ions or positively charged fragments from labile organometallic species. Once loaded, the metal ions can be converted to oxides, chalcogenides, or the zerovalent metal through exposure to the appropriate second reagent. Figure 1 illustrates the particular example of loading the nanoreactors with cadmium followed by H_2S treatment to form CdS clusters, confined within the block copolymer domain. The receptor sites are regenerated leaving the nanoreactors ready for further loading/reaction cycles with the same or different reagents. In principle, doped semiconductor clusters, metallic alloys, side-by-side cluster populations, and core–shell structures can be made via this universal nanoreactor approach, as reviewed by Ciebien *et al.* (1998).

Because the spherical domains of a block copolymer are generally monodisperse in size, the number of sequestering groups in every domain should be nearly identical, and this should lead to a predictable, monodisperse cluster size distribution. Cluster size should be proportional to domain size, both being dictated by the chain length of the block sequence containing the reagent-sequestering groups. Although Chan *et al.* (1992c) have managed to generate single, monodisperse silver clusters in certain circumstances, the more general case is shown in Fig. 2 where multiple clusters appear. Both thermodynamic and kinetic limitations on cluster growth in polymer films have been addressed by Kane *et al.* (1999). The block copolymer nanoreactor scheme appears to be more useful for spatial control of functional materials than for size control of inorganic clusters synthesized in this way.

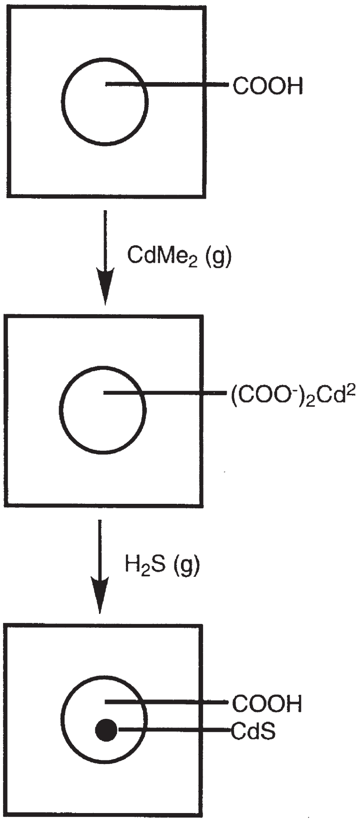


Figure 1
Schematic representation of the block copolymer nanoreactor approach to inorganic cluster synthesis.

2. Preloaded Nanoreactor Based on Organometallic Monomers

In a second approach, organometallic monomers are used to form one of the block sequences of the copolymer (Cummins *et al.* 1991). Spontaneous ordering of the morphology then amounts to the formation of preloaded nanoreactors ready for cluster-forming chemistry. This method eliminates the loading step of Fig. 1, which may be slow and possibly lead to nonuniform and nonselective sequestering of the reagents, but does require a new organometallic monomer for each target application. Figure 3 shows the appearance of palladium clusters in a reduced copolymer film in which one of the blocks comprises organopalladium repeat units.

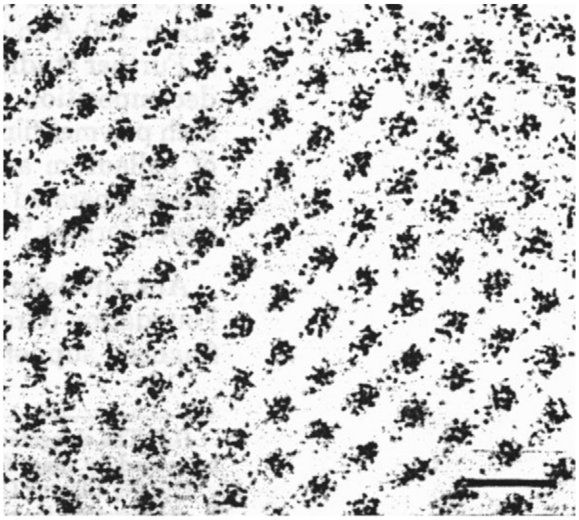


Figure 2
Transmission electron micrograph showing multiple gold nanoclusters in each domain. Scale bar = 500 Å (reproduced by permission of the American Chemical Society from *Chem. Mater.*, 1992, **4**, 24–7).

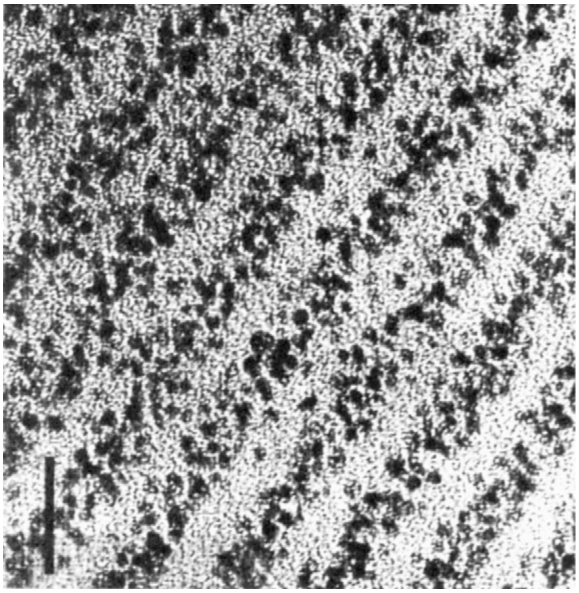


Figure 3
Palladium nanoclusters synthesized in the lamellar morphology of a diblock copolymer with one block consisting of norbornene-derived palladium complexes. Scale bar = 250 Å (reproduced by permission of the American Chemical Society from *Chem. Mater.*, 1992, **4**, 885–94).

3. Self-assembling Block Copolymer Morphologies in the Presence of Preformed Inorganic Clusters

Monodisperse stabilized inorganic clusters (Stucky 1991, Murray *et al.* 1993) can be stored indefinitely for later incorporation into polymer films. If there are groups in one of the blocks which are capable of displacing stabilizing species on the surfaces of the nanoclusters, then the clusters can be selectively conveyed into this block domain during solvent processing. Nanocomposites formed in this way are shown in Fig. 4.

4. Nanoporous Block Copolymer Templates

Selective degradation of one domain and subsequent removal of the degradation products creates a set of nanochannels in which functional inorganic materials may be grown. In one example, Hashimoto *et al.* (1997) selectively ozone-degraded the 1,4-polyisoprene

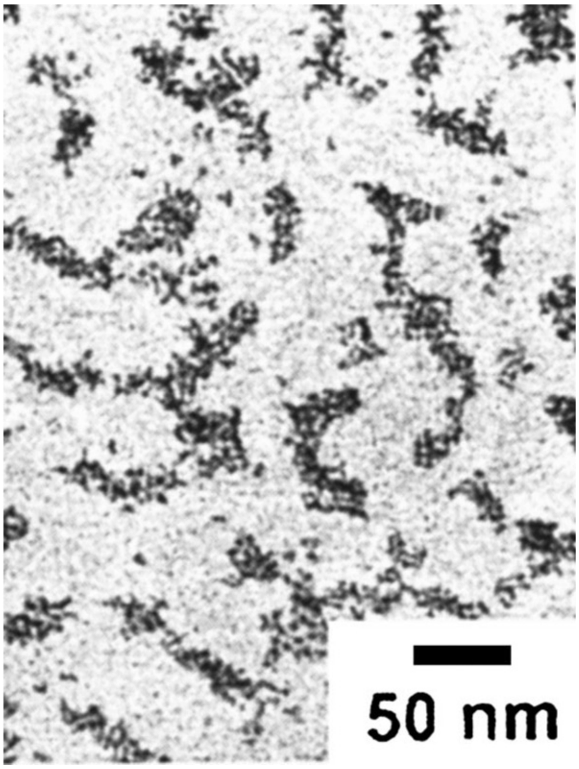


Figure 4
Cadmium selenide clusters selectively sequestered into domains containing an alkylphosphine moiety on the block repeat units (reproduced by permission of the American Chemical Society from *Macromolecules*, 1997, **30**, 8433–9).

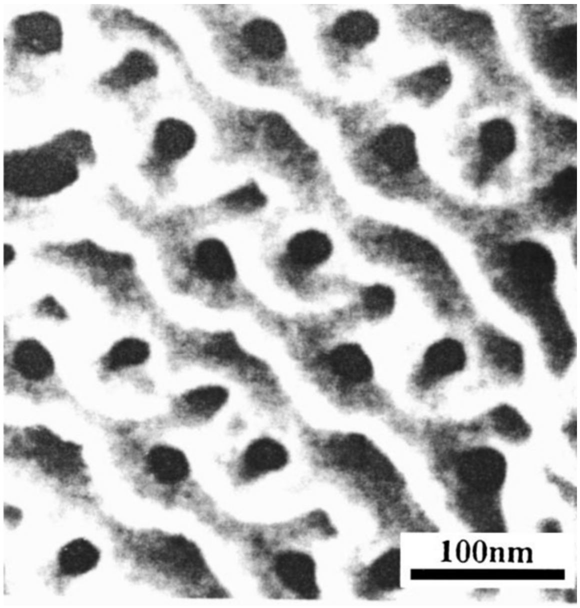


Figure 5
Bicontinuous nanochannel template suitable for synthesis of metal coating and semiconductor clusters (reproduced by permission of the American Chemical Society from *Langmuir*, 1997, **13**, 6869–72).

domains of the bicontinuous gyroid morphology of a polystyrene–polyisoprene diblock copolymer. The resulting nanoporous structure (Fig. 5) was suitable for metal plating with palladium, nickel, and gold. Transmission electron microscopy indicated that the plating process did not block the porosity of the final material, thereby rendering it suitable for subsequent gas–solid catalytic reactions.

For further details see *Block Copolymers as Precursors to Porous Materials** and *Block Copolymer Templates for Nanolithography*$^{\#}$.

5. *Applications*

The utility of the block copolymer template approach lies in the spatial control and uniformity of dispersion of the functional material in a thin film or bulk specimen. Most applications of these nanocomposites depend on the selection of the clusters or other functional materials to be spatially sequestered or synthesized within the morphology. For example, if permanent information storage is desired, magnetic clusters must be large enough and possess the necessary magnetic anisotropy to retain the imposed magnetization indefinitely. Films containing very small superparamagnetic clusters have potential for use as optically transparent, nonerasable magnetic watermarks. Similarly, electro-optical properties of the nanocomposite films are largely dictated by the properties of clusters embedded in the films. It is possible to envision applications that capitalize on both the ordered morphologies of the block copolymer template and its particular chemical composition, in addition to the functional characteristics of the sequestered clusters.

Catalytic membranes could be fashioned to exploit the selective gas transport characteristics of the domains of a diblock copolymer containing catalytically active clusters, facilitating both chemical reaction and product separation. Photovoltaic devices that rely on separate channels for transport of holes and electrons away from an optically excited site could be fashioned from ABC block copolymers which organize into an ordered morphology with semiconductor clusters at the interfaces between alternating domains comprising a hole-transport polymer and an election-transport polymer.

These examples suggest that block copolymer templates may offer certain advantages for producing functional materials. Independent control over the physical and chemical nature of the constituent blocks, and the properties of the sequestered (or *in situ* synthesized) functional material, provides a wide parameter space for the materials specialist seeking optimized performance in a variety of applications.

Bibliography

Chan Y N C, Schrock R R, Cohen R E 1992a Synthesis of palladium and platinum nanoclusters within microphase-separated diblock copolymers. *Chem. Mater.* **4**, 885–94

Chan Y N C, Schrock R R, Cohen R E 1992b Synthesis of silver and gold nanoclusters within microphase-separated diblock copolymers. *Chem. Mater.* **4**, 24–7

Chan Y N C, Schrock R R, Cohen R E 1992c Synthesis of single silver nanoclusters within spherical microdomains in block copolymer films. *J. Am. Chem. Soc.* **114**, 7295–6

Ciebien J F, Clay R T, Sohn B-H, Cohen R E 1998 Brief review of metal nanoclusters in block copolymer films. *New J. Chem.* **22**, 685–91

Cummins C C, Beachy M D, Schrock R R, Vale M G, Sankaran V, Cohen R E 1991 Synthesis of norbornenes containing tin(II), tin(IV), lead(II) and zinc(II) and their polymerization to give microphase separated block copolymers. *Chem. Mater.* **3**, 1153–63

Fogg D E, Radzilowski L H, Dabbousi B O, Schrock R R, Thomas E L, Bawendi M G 1997 Fabrication of quantum dot-polymer composites: semiconductor nanoclusters in dual-function polymer matrices with electron-transporting and cluster-passivating properties. *Macromolecules* **30**, 8433–9

Hashimoto T, Tsutsumi K, Funaki Y 1997 Nanoprocessing based on bicontinuous microdomains of block copolymers: nanochannels loaded with metals. *Langmuir* **13**, 6869–72

Kane R S, Cohen R E, Silbey R J 1999 Semiconductor nanocluster growth in polymer films. *Langmuir* **15**, 39–43

Martin C R 1994 Nanomaterials: a membrane-based synthetic approach. *Science* **266**, 1961

Murray C B, Norris D J, Bawendi M G 1993 Synthesis and characterization of nearly monodisperse CdE (E = S, Se, Te) semiconductor nanocrystallites. *J. Am. Chem. Soc.* **115**, 8706–15
Stucky G D 1991 *Naval Res. Rev.* **3**, 28

R. E. Cohen
Massachusetts Institute of Technology, Cambridge
Massachusetts, USA

Bone and Natural Composites: Properties

Nearly all biological skeletal materials are composites. Two composites will be dealt with that are particularly well-characterized anatomically: bone and insect cuticle. Unfortunately, little is known about the mechanical properties of insect cuticle.

The concept of "helicoids" will occur frequently. A helicoid is a multidirectional plywood. It consists of linear elements arranged in sheets. The linear elements are all oriented in the same direction in each sheet. The orientation changes between sheets (typically by between 10° and 20°), and the sense of rotation is constant between sheets. This description comes from Neville (1993), which should be referred to for a comprehensive discussion of helicoids and similar structures in nature.

1. *Bone: Structure*

Bone is a composite with three main constituents: the fibrous protein collagen; the mineral, carbonated apatite, close to dahllite in structure; and water. The relative weight percent values of the three components in "ordinary" bone, such as the long bone of mammals, is about 65%, 25%, and 10% for the collagen, mineral, and water, respectively. Most bone of birds and mammals is "lamellar" bone. This bone type has a hierarchical structure. At the lowest level are collagen fibrils infiltrated by and surrounded by mineral crystals. These fibrils tend to be organized in planar sheets, with a preferred orientation, and this preferred orientation may be more or less constant over many tens of micrometers. At the next level these sheets are packed together into repeating lamellae, about 4 μm thick. These lamellae may have most fibrils oriented in one direction, or they may change progressively through the thickness of the lamella in a helicoidal arrangement. At the next level the lamellae are arranged either in rather continuous blocks, many hundreds of micrometers in size, or may, as the result of the remodeling that occurs continually in bone, be arranged in tubular structures (secondary osteons, often called Haversian systems), of the order of 150 μm in diameter with the concentric lamellae surrounding axial blood vessel(s). This lamellar bone contains entombed bone cells, which connect with each other and with the nearest blood channels via tiny channels, about 200nm in diameter. Finally, these relatively large lumps of bone may form "compact" bone, containing no cavities obvious to the naked eye, or "cancellous" bone, formed of many interconnected struts or sheets, whose characteristic thickness is about 200mm. This is obviously porous to the naked eye. The struts are surrounded by marrow. The volume fraction of this cancellous bone is rarely more than 30%.

This description is of lamellar bone, of which Weiner *et al.* (1999) give a good review. Bone exists in another, much less well-structured version: woven bone. This is more highly mineralized and isotropic or, rather, random in its structure, and is almost certainly weaker. It has the advantage from the animal's point of view that it can be laid down much faster than lamellar bone. Often woven and lamellar bone are incorporated in a multiple sandwich into the final structure. This is called "fibrolamellar" bone.

2. *Bone: Mechanical Properties*

2.1 *Elastic Properties*

Table 1 gives characteristic values for "ordinary" bone as determined both ultrasonically and mechanically.

2.2 *Strength Properties*

There is considerable difference in the strength and ultimate strain when bone is loaded in its different directions relative to the long axis of the whole bone (Table 2). Weiner *et al.* (1999) give more information about off-axis behavior.

2.3 *Fracture Mechanics Properties*

Some of the criteria for the successful application of fracture mechanics analyses to bone are not fulfilled. In particular, bone frequently undergoes considerable damage well away from the point where the fracture will eventually occur, even if this fracture is induced by a sharp notch. A few values, which seem to be representative of values in the literature, are given in Table 3.

Note that in life the direction in which a dangerous crack is likely to travel is transversely, and so the longitudinally traveling cracks are not really a useful measure of the service toughness of the bone.

2.4 *Fatigue*

Bone often fails after cyclic loading. The results of several studies are shown in Fig. 1. This is a conventional "*S–N*" diagram showing the relationship

Table 1
Elastic moduli for bone (gigapascals).

	Ashman *et al.* (1984)		Reilly and Burstein (1975)				
	Canine	Human	Human		Bovine		
	(ultrasound)	Haversian (ultrasound)	Haversian (mechanical)		Haversian (mechanical)		Fibrolamellar (mechanical)
			Tension	Comp	Tension	Comp	Tension
E_1	12.8	12.0	12.8	11.7	10.4	10.1	11.0
E_2	15.6	13.4	12.8	11.7	10.4	10.1	11.0
E_3	20.1	20.0	17.7	18.2	23.1	22.3	26.5
G_{12}	4.7	4.5					
G_{13}	5.7	5.6	3.3		3.6		5.1
G_{23}	6.7	6.2	3.3		3.6		5.1
ν_{12}	0.28	0.38	0.53	0.63	0.51	0.51	0.63
ν_{13}	0.29	0.22					
ν_{23}	0.26	0.24					
ν_{21}	0.37	0.42	0.53	0.63	0.51	0.51	0.63
ν_{31}	0.45	0.37	0.41	0.38	0.29	0.40	0.41
ν_{32}	0.34	0.35	0.41	0.38	0.29	0.40	0.41

These values were determined by Reilly and Burstein (1975) and Ashman *et al.* (1984). These are old papers, but more recent results are very similar. The subscripts "1", "2," and "3" refer to the radial, tangential and longitudinal directions relative to the long axis of the bone. Reilly and Burstein did not test in all directions, and assumed that the bones were transversely isotropic. Shear values (G) are included under the heading of "tension" merely for convenience. They assumed values and G_{13} and G_{23} to be equal, and also pairs of Poisson's ratio (ν): ν_{12}, ν_{21} and ν_{31}, ν_{32}.

Table 2
Strength properties of bone (megapascals).

	Human		Bovine					
	Haversian		Haversian			Fibrolamellar		
	Longitudinal	Transverse	Longitudinal	Transverse	Radial	Longitudinal	Transverse	Radial
Tension								
Strength	133	53	150	54	39	167	55	30
Yield stress	114		141			156		
Ultimate strain	0.038	0.007	0.020	0.007	0.007	0.033	0.007	0.002
Compression								
Strength	205	131	272	171	190			
Ultimate strain	0.019	0.050		0.042	0.072			
Shear								
Strength	67		70			64		

The values were taken directly, or from information in Reilly and Burstein (1975), and Reilly *et al.* (1974). "Longitudinal," "transverse," and "radial": specimens tested longitudinally, transversely, or radially with respect to the long axis of the bone. Shear strengths are placed arbitrarily in the columns, since shear refers to two axes at the same time.

between the imposed stress and the number of cycles required to produce failure. Fatigue loading can be considered (except in the case of the rotating cantilever) to be a continually interrupted creep rupture stress. By loading at different frequencies it is theoretically possible to distinguish the effects of time and cycles. The evidence suggests that in the tension cycle, numbers are unimportant but that in compression the situation is more complex (Caler and Carter 1989).

2.5 The Range of Properties

The properties referred to previously are for "ordinary" bone, mainly long bones of bovines and humans.

Table 3
Some fracture mechanics properties of bone.

Species	Direction	K_c (MN m$^{-3/2}$)	G_c (kJ m^{-2})	Test	Rate
Bovine	Longitudinal	3.2		SENT	Slow
Bovine	Longitudinal	5.0		SENT	Fast
Bovine	Transverse	4.5		SENT	Slow
Bovine	Transverse	7.7		SENT	Fast
Bovine	Longitudinal	5.3	1.2	CT	Slow
Bovine	Longitudinal	6.3	2.9	CT	Fast
Bovine	Transverse	6.5		CT	Slow
Human	Longitudinal	4.0	0.5	CT	Slow
Antler	Transverse	5.4		SENT	Slow

"Direction," direction of crack travel relative to the long axis of the bone; "SENT," single edge notch, tension; "CT," compact tension.

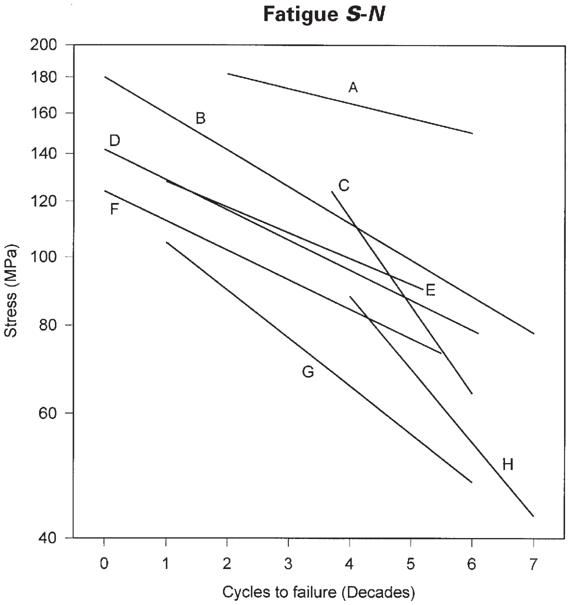


Figure 1
A set of S–N curves for bone. The separate lines are the lines representing the equations produced by regression analysis. The lines are not extrapolated beyond the ends of the original data. In all the cases the distributions were reasonably linear. The abscissa is measured in powers of 10. Taken from various experiments, with different loading conditions. A: human, four-point bending, room temperature, 2 Hz. B: human, compression, room temperature, 30 Hz. C: bovine, rotating cantilever, room temperature, 30 Hz. D: bovine, tension, room temperature, 2 Hz. E: human (27 years old) tension, room temperature, 2 Hz. F: red deer antler, tension, room temperature, 2 Hz. G: human (69 years old) tension, 37 °C, 2 Hz. H: human, rotating cantilever, room temperature, 70Hz.

However, the properties of bone can range considerably outside the values so far described, mainly as a result of differences in the mineral content, and in the degree of anisotropy of the tissue. Table 4 gives values (from Currey 1998) for Young's modulus, tensile strength, ultimate tensile strain, and work under the stress–strain curve for various bony tissues, which all had essentially the same shape, so the results are truly comparable. They are listed in order of increasing mineral content (milligrams of calcium per gram of dried, defatted bone). Note particularly how the ultimate strain decreases with increase in mineral content. In general these differences in mechanical properties seem to accord with the requirements of the tissue, for instance deers' antlers need to be good at resisting impact while, for acoustic reasons, the ear bone of the whale needs to be very stiff and, being protected inside the skull, its poor strength properties do not matter. Figure 2, derived mainly from this data set, shows the apparent impossibility of marrying a high Young's modulus and a high ultimate strain (which is a good indicator of toughness).

2.6 *Cancellous Bone*

(*a*) *Structure properties.* Cancellous bone, found mainly at the ends of long bones, is made of interconnecting plates and rods that rarely achieve a volume fraction of greater than 30%. Its mechanical properties are very varied, depending on the amount of bone present, and its anatomical anisotropy. The best general treatment is by Rice *et al.* (1988). Both modulus and strength increase approximately as the square of the volume fraction of bone, and the Young's modulus of elasticity is about 100 times the compressive strength.

Table 4
Some mechanical properties of various bone tissues.

Species and tissue	Calcium content (mg g^{-1})	E (GPa)	s_{ult} (MPa)	e_{ult}	W (MJ m^{-3})
Red deer, immature antler	212	11.5	283	0.106	2060
Red deer, mature antler	214	8.7	165	0.098	1167
Sarus crane, ossified tendon	224	17.0	218	0.062	277
Polar bear (3 months), femur	235	6.7	85	0.044	327
Narwhal, tusk cement	235	5.3	84	0.060	300
Narwhal, tusk dentine	240	10.3	120	0.037	370
Sarus crane, tarsometatarsus	240	23.1	218	0.018	200
Axis deer (term fetus), tibia	242	13.2	124	0.029	190
Walrus, humerus	245	14.2	105	0.026	139
Fallow deer, radius	249	25.0	219	0.020	213
Human, adult, femur	249	16.7	166	0.029	274
Bovine, tibia	252	20.4	164	0.018	198
Leopard, femur	254	21.5	215	0.034	342
Brown bear, femur	255	16.9	157	0.032	234
Donkey, radius	255	16.2	127	0.024	210
Sarus crane, tibiotarsus	256	22.6	244	0.032	406
Flamingo, tibiotarsus	257	28.2	205	0.015	185
Roe deer, femur	259	17.2	155	0.014	132
Polar bear (3 years), femur	259	17.7	153	0.026	204
Fallow deer, tibia	259	25.2	167	0.009	88
King penguin, humerus	266	22.0	154	0.007	55
King penguin, ulna	265	22.2	187	0.010	94
King penguin, radius	266	22.6	186	0.010	80
Wallaby, tibia	268	25.4	187	0.008	86
Polar bear (7 years), femur	268	22.2	161	0.020	149
Bovine, femur	270	24.5	142	0.006	35
Axis deer, tibia	273	31.6	221	0.019	100
Wallaby, femur	274	22.0	177	0.012	146
Horse, femur	277	23.7	157	0.010	84
Fin whale, ear bone	315	34.0	27	0.0018	2

(*b*) *Material properties.* There seems to be no particular reason why the mechanical properties of cancellous bone *material* should differ much from those of compact bone. However, there have been studies suggesting that cancellous bone material is less stiff and, therefore, probably weaker than compact bone material.

Cancellous bone trabeculae are small, ragged, and thin, and difficult to test, and in testing very small specimens all sorts of size artifacts appear. Various methods have been used, including back-calculation from finite element models.

Some representative results are shown in Table 5. The consensus is that the Young's modulus for the *material* of cancellous bone is lower than that of neighboring compact bone, but it is not clear by how much. In general, the more recent studies have suggested the higher values for Young's modulus. Values that are less than about 6 GPa should be treated with skepticism because values as great as this have been obtained for *whole blocks* of cancellous bone, completely discounting the considerable effects of the porosity itself.

If cancellous bone material has considerably lower mechanical properties than that of compact bone there is the question as to why this should be so. There are at least two possible reasons. One is that the mineral content of cancellous bone is less than that of neighboring compact bone. There is some evidence for this although the difference is not large. Another possible reason is that although the alignment of the lamellae and the general grain of cancellous bone are roughly along the line of the trabeculae, in detail it is not as well organized as it is in compact bone. Even this argument is somewhat dubious because compact bone loaded at about 90° to its predominant orientation, the worst possible orientation, has a Young's modulus about half that of bone loaded along its grain. Cancellous bone is much better oriented along the trabeculae than compact bone loaded at right angles to its grain.

3. Insect Cuticle

The skeleton of insects is an "exoskeleton" forming an almost complete covering to the body. To allow movement, stiff plates or tubes (sclerites) alternate with flexible "arthrodial membranes." However, the chemical compositions of the stiff and compliant regions are rather similar. The main fibrous component of cuticle is chitin, a linear polysaccharide of poly-*N*-acetylglucosamine. Groups of about 20 molecules pack together, forming rods of very high aspect ratio, about 4nm in diameter. The matrix of insect cuticle is almost entirely protein, and there are many different component proteins. The protein is probably firmly attached to the chitin, by covalent bonds, but this is not certain (Andersen *et al.* 1996). The detailed arrangement varies somewhat, but the water bug *Hydrocirus* is representative. Its chitin fibrils are about 4.5nm in diameter, arranged in layers, the thickness of the layers typically being about 6nm. They are arranged either helicoidally, in which case the preferred direction changes through 180° in about 25 layers over a distance of 160nm, or unidirectionally. A remarkable (and unexplained) feature of many cuticles is that the cuticle laid down at night is arranged helicoidally but the cuticle laid down during the day is unidirectional. Sometimes the plies seem orthogonally arranged, but close examination always shows that there are some helicoidal layers leading from one orientation to the next. Naturally, with this range of possible arrangements, it is possible for almost any multilayer composite arrangement to be produced.

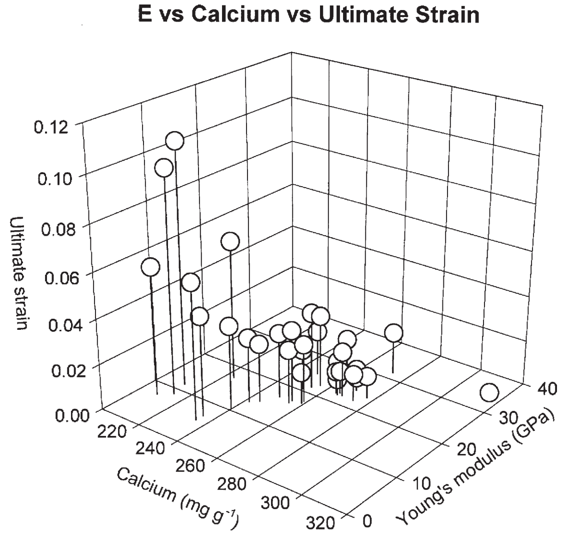


Figure 2
Three-dimensional diagram showing the relationship between mineral content, Young's modulus, and ultimate strain of various bony tissues. The data set is nearly the same as given in Table 4.

The mechanical properties of the cuticle can be profoundly affected by the extent to which the protein matrix is cross-linked or "sclerotized." The process is complex, but essentially involves the oxidation of *N*-acetyldopamine and *N*-β-alanyldopamine to the corresponding *o*-quinones. These products may then be further modified, and become linked to histidine and lysine residues in the cuticular proteins, and may

Table 5
Estimates of Young's modulus of cancellous bone material.

Bone	Method	Young's modulus (GPa)
Human proximal femur	Buckling	11.4
Bovine femur	Tension	0.8
Human tibia	Three-point bending	3.2
Bovine femur	Ultrasonic	10.9
Human femur	Ultrasonic	12.7
Human ilium	Three-point bending	4.6
Human L3 vertebra	Structural analysis	3.8
Human tibia	Four-point bending	5.4
Human tibia, trabeculae	Ultrasonic	14.8
Human tibia, cortical	Ultrasonic	20.7
Bovine proximal femur	Extrapolation from microhardness	15
Human femur, trabeculae	Ultrasonic (wet)	17.5
Human femur, cortical	Ultrasonic (wet)	17.3
Human femur, trabeculae	Nanoindentation (dry)	18.1
Human femur, cortical	Nanoindentation (dry)	20.0
Human tibia	Three-dimensional FEA	5.9
Human tibia	Three-dimensional FEA	14.5

Table 6
Some mechanical properties of insect cuticle.

Animal	Cuticle type	Young's modulus (Pa)
Locusta (locust)	Female intersegmental membrane in laying mode	10^3
Schistocerca (locust)	Tendon (along length)	1.1×10^{10}
	Tendon (normal to length)	1.5×10^8
	Tibia cuticle	1.7×10^{10}
Calliphora (blowfly)	Unsclerotized pupal skin	7.3×10^7
	Sclerotized pupal skin	2.4×10^8
Phormia (hoverfly)	Wing	6.1×10^9
Rhodnius (blood-sucking bug)	Abdominal cuticle (unplasticized)	2.4×10^8
	Abdominal cuticle (plasticized)	2.5×10^6

Animal	Cuticle type	Strength(Pa)
Schistocerca (locust)	Tibia cuticle	9.5×10^7
	Tendon	6×10^8

also become attached to the chitin fibrils (Andersen *et al.* 1996). It is possible that the concomitant reduction in water content may also simultaneously affect the mechanical properties. Usually the cuticle of insects has an outer layer (the exocuticle) which is heavily sclerotized, and whose chitin is arranged helicoidally, and an inner layer (the endocuticle) which is not sclerotized, and whose chitin may be arranged helicoidally, or in various other modes.

Unlike bone, which comes in relatively large lumps, insect cuticle is often thin, highly curved, and, as in exo- as opposed to endocuticle, extremely inhomogeneous. These features make mechanical testing difficult, and in fact although many attempts to characterize mechanical properties were made in the 1970s and early 1980s, interest declined, and only revived at the end of the 1990s with the construction of microtesting machines.

Table 6 gives values for modulus and for strength. Most of these values were originally collated by Vincent (1980). The female locust has an extremely extensible, very low-modulus intersegmental membrane that allows her to deposit eggs deep in the soil. The mechanical anisotropy of the tendon is extremely high; this is because it is essentially a unidirectional composite. Applying composite theory to this tendon allowed Ker (1977)) to estimate the moduli of the chitin and the protein components. He obtained values of 80 GPa and 0.12 GPa, respectively, a remarkably large difference between fibers and matrix. The blood-sucking bug *Rhodnius* is able to plasticize its abdominal cuticle reversibly, so that it is very compliant when a meal is being taken but becomes stiffer at other times. The very high value for tendon strength (loaded along its long axis) is probably due, again, to its being a very highly oriented composite. We do not know its strength in the orthogonal direction.

4. Conclusion

Although bone and cuticle have been recognized as composite materials for a long time, their structural complexity (bone) or small size and convoluted shape (cuticle) have prevented satisfactory analysis. It is to be hoped that, as micro-testing methods develop, this will change, and we will gain a better understanding of these fascinating natural composites.

See also: Cuticle#

Bibliography

Andersen S O, Peter M G, Roepstorff P 1996 Cuticular sclerotization in insects. *Comp. Biochem. Physiol.* **113B**, 689–705

Ashman R B, Cowin S C, Van Buskirk W C, Rice J C 1984 A continuous wave technique for the measurement of the elastic properties of cortical bone. *J. Biomech.* **17**, 349–61

Caler W E, Carter D R 1989 Bone creep—fatigue damage accumulation. *J. Biomech.* **22**, 625–35

Currey J D 1998 Mechanical properties of vertebrate hard tissues. *J. Eng. Med.* 399–411

Ker R F 1977 Some structural and mechanical properties of locust and beetle cuticle. D.Phil. thesis, University of Oxford

Neville A C 1993 *Biology of Fibrous Composites: Development Beyond the Cell Membrane*. Cambridge University Press, Cambridge

Reilly D T, Burstein A H 1975 The elastic and ultimate properties of compact bone tissue. *J. Biomech.* **8**, 393–405

Reilly D T, Burstein A H, Frankel V H 1974 The elastic modulus for bone. *J. Biomech.* **7**, 271–5

Rice J C, Cowin S C, Bowman J A 1988 On the dependence of elasticity and strength of cancellous bone on apparent density. *J. Biomech.* **21**, 155–68

Vincent J F V 1980 Insect cuticle: a paradigm for natural composites. In: Vincent J F V, Currey J D (eds.) The mechanical properties of biological materials. *Symp. Soc. Exp. Biol.* Cambridge University Press, Cambridge, Vol. 39, pp. 183–210

Weiner S, Traub W, Wagner H D 1999 Lamellar bone: structure–function relations. *J. Struct. Biol.* **126**, 241–55

J. D. Currey
University of York, UK

Bone Augmentation and Repair

Bone is a complex tissue that performs a variety of integral functions in the maintenance of many systems of the body. Bone is the primary reservoir of calcium, phosphate, and other important ions, and it acts to regulate the concentration of these minerals throughout the body. In addition, the marrow found within the bone generates red and white blood cells for immunoprotection and oxygenation of other tissues. Finally, bone provides structural support and protection and sites of muscle attachment for locomotion (Yaszemski *et al* 1996). Under normal conditions, bone has the ability to repair and regenerate itself without the formation of a scar. Under certain circumstances, however, this function of the body fails and intervention is required.

This article begins with an overview of the structural properties and natural repair mechanism of bone. It then focuses on the clinical methods and materials used for bone augmentation and repair. Finally, it addresses the development of synthetic, bioresorbable polymers for use as scaffolds for cell transplantation or conduits for guided tissue ingrowth.

1. Bone Physiology

Mature bone is a highly vacularized tissue composed of cells and an extracellular matrix (ECM). The cells include bone-forming cells (osteoblasts), bone-resorbing cells (osteoclasts), and bone-maintaining cells (osteocytes). The ECM is a composite material consisting of organic and inorganic components. The organic component, also called the osteoid, is secreted by the osteoblasts. The osteoid constitutes approximately 50% of bone by volume (25% by weight) and is composed of 90% Type I collagen and 10% amorphous ground substance (primarily glycoprotiens and glycosaminoglycans) (Bouvier 1989). The organic component gives bone its form and contributes to its ability to resist tension (Buckwalter *et al.* 1995). The inorganic component is a mineral phase composed of calcium crystals primarily in the form of hydroxyapatite ($[Ca_{10}(PO_4)]_6(OH)_2$). This component makes up 50% of bone by volume (75% by weight) (Bouvier 1989) and gives bone most of its stiffness and strength (Buckwalter *et al.* 1995).

When bone is initially laid down, it is structurally weak and unorganized (Yaszemski *et al.* 1996). Within days, this primary bone is remodeled to become lamellar bone. Lamellar bone has alignment of the collagen fibers and mineral phase. Mature lamellar bone exists as cortical (also called compact) or trabecular (cancellous, spongy) bone. Cortical bone is found on the outer surface of most bones. It is dense with respect to trabecular bone (apparent density equal to $1.8\,\mathrm{g\,cm^{-3}}$ for cortical vs. $0.1–1.0\,\mathrm{g\,cm^{-3}}$ for trabecular), and contains closely spaced groups of concentric lamellar bone rings called osteons (Yaszemski *et al.* 1996). Cortical bone functions mechanically in tension, compression, and torsion (Buckwalter *et al.* 1995). Trabecular bone occurs near the ends of long bones and in the center of small and flat bones. It is composed of an array of plates and rods that form an open cell foam (Yaszemski *et al.* 1996). Trabecular bone functions mainly in compression (Buckwalter *et al.* 1995).

Bone has the unique ability to remodel itself along lines of stress in response to injury or other changes in its mechanical environment (Bostrom and Mikos 1997). When a bone is fractured, the body creates new bone to connect the broken fragments, and then remodels the new bone to optimize its mechanical function in the particular region of the skeleton where the fracture has occurred.

This repair process begins immediately (Yaszemski *et al.* 1996). The broken blood vessels in the vicinity of the fracture fill the space in and around the fractured bone ends with a fracture hematoma. New blood vessels grow into the defect, and bone-forming cells migrate into the area. The cells (osteoblasts) secrete the osteoid, which will be mineralized and eventually become woven bone. Woven bone is isotropic, and will be remodeled by the sequential action of removal and redeposition along stress lines. Remodeling occurs continuously in the body, and results in the anisotropic structure of lamellar bone that is mechanically and structurally optimized for each area of the skeleton (Thomson *et al.* 1995).

As many as 5–10% of the 6.2 million fractures sustained by patients in the USA each year result in fracture nonunion and decreased mechanical support (Einhorn 1999). When this occurs, an attempt must be made to induce bone regeneration for skeletal repair or reconstruction. The surgeon strives to make local skeletal conditions mimic those that the body would expect to encounter during normal fracture repair. In essence, it is attempted to "trick" the body into initiating the fracture repair cascade (Thomson *et al.* 1995).

2. *Techniques*

The "gold standard" for bone augmentation and repair is the trabecular bone autograft. Transplantation of autologous tissue involves the harvesting of bone from a distant location and implanting it into the defect site. A similar approach uses bone harvested from cadavers (allogenous tissue). Both techniques are aimed at the ultimate restoration of bone tissue. Transplantation provides a scaffold similar to the surrounding bone upon which bone can invade, lay down extracellular matrix, and remodel. Although these treatments are often successful, they have their associated problems and limitations. Autologous tissue is limited in supply, is often not available in the shape needed, and its relocation leaves a secondary defect site. Allogenous tissue runs the risk of pathogen transfer and immune system rejection of the foreign material. Both transplantation techniques injure local vessels and leave bleeding host bone at the defect site that will activate the coagulation cascade and may lead to necrosis of the grafted tissue (Yaszemski *et al.* 1996).

Poly(methyl methacrylate) (PMMA) is a nondegradable bone cement that is widely used to fill bone defects or provide fixation of prostheses to bone. PMMA is a biologically inert material which acts as a permanent barrier preventing fracture healing or integration with host bone. A study of 22 commercially available plain PMMA-based bone cements reports compressive strengths in the range 75–115 MPa and moduli of elasticity of 1700–3100 MPa (Kuhn 2000). Because these values are much greater than those of trabecular bone (5–10 MPa and 50–100 MPa, respectively) (Temenoff *et al.* 2000) the use of PMMA bone cements can lead to transfer of load to the material (stress shielding) and insufficient mechanical stimulation to adjacent tissue. Lack of proper mechanical loading induces bone resorption around the implant and fatigue failure of the implant (Bostrom and Mikos 1997). In addition, a number of PMMA cements have been shown to shrink more than 5% during polymerization (Kuhn 2000), which may lead to micromotion of the implant and an inflammatory response within the surrounding tissue.

3. *Tissue Engineering*

The limitations associated with regeneration techniques have inspired researchers to search for other methods to repair skeletal defects. New strategies to engineer bone have focused on the use of natural or synthetic materials as scaffolds for cell transplantation or conduits for bone ingrowth. Various types of scaffolds have been tested, but synthetic bioresorbable polymers that degrade by hydrolysis into natural metabolites of the body remain the most attractive. These materials are easily mass-produced, and changing the composition of the material can precisely modulate their physical, chemical, mechanical, and degradative properties. Scaffolds composed of copolymers or polymer blends can be manufactured to create materials with properties that are advantageous over homopolymers for specific applications. Biodegradable polymers offer the possibility to completely replace osseous tissue and thus overcome problems such as infection and device dislocation associated with permanent implants (Temenoff *et al.* 2000).

3.1 Scaffold Properties

Synthetic polymeric scaffolds must possess unique properties in order to optimally enhance bone formation (see Table 1). The primary role of a scaffold is to provide a temporary substrate to which cells can adhere (Thomson *et al.* 1995). How cells attach to the scaffold can affect their behavior: strong cell adhesion and spreading are associated with proliferation, while a rounded cell shape is indicative of cell-specific function (Temenoff *et al.* 2000). A scaffold must therefore act as a substrate to promote cell adhesion and maintenance of function without hindering proliferation. The effectiveness of a material in achieving this goal is dependent mainly on surface chemistry, which determines the interactions between cell and substrate (Thomson *et al.* 1995). In addition, the spatial relationship of cells within their ECM can be important for continuance of function. Therefore, three-dimensional scaffolds can assist in the organization of cell growth and the formation of ECM (Thomson *et al.* 1995).

Porosity, pore size, and pore structure are important parameters in ensuring adequate room for tissue regeneration as well as nutrient delivery to cells within a three-dimensional construct. Highly porous scaffolds with large void volume and high surface area to volume ratio are desirable to maximize space for cell seeding, attachment, growth, ECM production, and vascularization. To attain the high surface area per unit volume needed for anchorage-dependent cells, smaller pore diameters are preferred as long as the pore size is greater than the diameter of a cell in suspension (typically 10 μm). However, larger pores are required for cell migration, growth, and ECM production. Experiments suggest that a pore diameter of 200–400 μm is optimal for bone ingrowth (Temenoff *et al.* 2000). An interconnected pore network enhances diffusion within the scaffold, improving nutrient supply and waste removal, thus increasing viability of cells beyond the edges of the construct.

Initial mechanical properties of the polymer scaffold should be almost identical to the surrounding bone. If the scaffold is much stronger than bone, stress shielding occurs (Bostrom and Mikos 1997). On a microscopic scale, the local stiffness of the

Table 1
Scaffold design parameters.

Porosity	Maximum without compromising mechanical properties
Pore size	200–400 μm
Pore structure	Interconnected
Mechanical properties	
Trabecular bone	Tension and compression: strength, 5–10 MPa; modulus, 50–100 MPa
Cortical bone[a]	Tension: strength, 80–150 MPa; modulus, 17–20 GPa Compression: strength, 130–225 MPa; modulus, 17–20 MPa Shear: longitudinal, 54–70 MPa; radial, 3.3 GPa
Degradative properties	
Degradation time[b]	Less than 6 months
Degradation mechanism	Bulk or surface erosion
Time to reach gel point[c]	~10 min
Biocompatibility	No harmful chronic inflammation
Sterilizability	Sterilizable without altering material properties

Source: Temenoff *et al.* (2000).
[a] Reported values for the strength and modulus of cortical bone were measured parallel to the long axis of the osteon. [b] Degradative time in this instance refers to the time to loss of scaffold mechanical properties (time at which the scaffold no longer shares in weight bearing). [c] Time to reach gel point is for injectable, *in situ* cross-linkable systems.

polymer may affect the mechanical tension generated by the cytoskeleton of a cell, which can control cell shape and therefore function (Temenoff *et al.* 2000). As well as providing proper support in the early stages of healing, graded load transfer is needed later in the process for creation of replacement tissue that is identical to the original. The rate of scaffold degradation, therefore, must balance the need to maintain structural support for cellular proliferation and ECM secretion with the need for polymer disintegration to leave room for new tissue ingrowth.

The mechanism by which the scaffold degrades and the degradation by-products should both be considered when developing a polymeric scaffold. Materials that undergo surface erosion degradation release products gradually. The continued renewal of the exterior of this material renders cell attachment and culture on these materials problematic. Materials that degrade through bulk erosion release their products only when the molecular weight of the polymer reaches a critical value. Past the critical point, mechanical properties are difficult to maintain. Additionally, when the degradation products are acidic, a burst effect may cause a substantial drop in the local pH, which in turn may cause tissue necrosis or inflammation. Micromotion at the implant site or microparticles formed during the degradation process may elicit an inflammatory response. For example, it has been demonstrated that polymer microparticles suppress initial rat marrow stromal osteoblast proliferation and differentiation *in vitro* (Wake *et al.* 1998).

Any material proposed for implantation must meet certain general criteria. The material must be biocompatible. In other words, it must not evoke an extreme adverse inflammatory or immunological response. Synthesis, processing, or sterilization techniques can affect certain factors that determine biocompatibility, including polymer chemical and physical structure and surface morphology. Residual chemicals involved in these processes may leach out of the construct and harm the surrounding native tissue and the newly formed tissue within the construct. Finally, to create a truly viable alternative for bone repair, a polymer must be able to be easily sterilized, be able to be manufactured reproducibly on a large scale, have a long shelf life, and be easy for surgeons to handle in an operating room environment.

3.2 Material Selection

Scaffold design parameters may be optimized through the selection of the proper biomaterial. As with the design parameters, the choice of the material depends greatly on its intended application. The poly(α-hydroxy acid) family, which includes poly(lactic acid) (PLA), poly(glycolic acid), and poly(lactic-*co*-glycolic acid) (PLGA) copolymers, has been extensively investigated for use as prefabricated scaffolds. These polymers are particularly attractive as they are already approved by the US Food and Drug Administration for certain clinical applications. The degradation products, lactic and/or glycolic acid, are found naturally in the body and are therefore processed through normal metabolic pathways (Yaszemski and Yasko 1998). However, their buildup at the implant site can lower the local pH and alter the rate of degradation.

Other polymers have been explored for their use as preformed scaffolds for bone regeneration including polyanhydrides and polyphosphazenes. Unlike polyesters that degrade by bulk erosion, polyanhydrides degrade in a controlled fashion by surface erosion. Although this type of degradation may result in a more controlled release of by-products, these materials often display poor mechanical properties. Polyanhydride-*co*-imides have been developed to augment the strength of the polyanhydride scaffold, but imides are not biodegradable (Behravesh *et al.* 1999). Moreover, surface-eroding polymers do not facilitate cell adhesion. Polyphosphazenes have a phosphorus–nitrogen backbone, and are rendered hydrolytically degradable by the addition of imidazoyle or amino acid alkyl ester side chains. The nature of the side chain and the percent substitution can affect the degradation rate and cell attachment (Yaszemski and Yasko 1998).

Poly(propylene fumarate) (PPF) is another material extensively studied for use as a tissue guidance scaffold. PPF is an unsaturated linear polyester that can be cross-linked through its fumarate double bond. PPF degrades by hydrolysis into products that occur naturally in the body (He *et al.* 2001). In an attempt to encourage host cell attachment and migration into the scaffold, PPF has been modified with GRGD (glycine–arginine–glycine–apartic acid) peptide sequences (Jo *et al.* 2000). Both prefabricated and injectable formulations of PPF have been investigated.

Injectable polymers are valuable because they can fill irregularly shaped defects with minimal surgical intervention. In addition to the design parameters discussed above, injectable systems need to meet the additional criteria of cross-linking *in situ* in a timely fashion without detrimental effects to the surrounding tissue. All reagents, including the initial polymer/monomer formulations, initiators, cross-linking agents, porogens, and reaction products, must be nontoxic. Furthermore, reaction parameters such as heat release and pH change must be minimized.

When cross-linked *in situ*, PPF is injected into the defect with any one of a number of cross-linking agents (e.g., *N*-vinyl pyrrolidinone (Peter *et al.* 1999), poly(ethylene glycol)-dimethacrylate (He *et al.* 2000), or PPF-diacrylate (He *et al.* 2001)), the initiator benzoyl peroxide, and the leachable porogen sodium chloride (NaCl). Many of the scaffold parameters can be controlled through manipulation of the method of PPF synthesis, cross-linking agent used, and ratios of initiator and cross-linking agent to PPF. Injectable formulations of PPF have been produced that reach gel point between 5 min and 10 min and reach a maximum cross-linking temperature only 10 °C higher than the normal physiological temperature of 37 °C. In addition, they have the desired compressive strength and modulus, both initially and during the course of degradation, for consideration for trabecular bone augmentation (Peter *et al.* 1997, 1998, 1999).

3.3 *Scaffold Applications*

One of the most promising techniques using bioresorbable polymer scaffolds for skeletal regeneration involves the transplantation of autologous cells. Pre-osteoblasts obtained from the patient can be harvested from a remote location, expanded and differentiated in culture, and seeded onto the appropriate scaffold (Crane *et al.* 1995). After further culture of the cells within the scaffold, the construct is transplanted to the wound site. This method is the most widely used in tissue engineering and allows the cells of the patient to fabricate an ECM with the appropriate structural properties and signaling molecules under controlled conditions in culture. All of the prefabricated scaffold materials discussed above have been investigated for this strategy, while most of the research has focused on the use of poly(α-hydroxy acid)s. Osteoblast/PGA constructs implanted subcutaneously in nude mice and in full thickness rabbit calvarial defects were shown to form new bone after 20 weeks (Vacanti *et al.* 1993) and 12 weeks (Breitbart *et al.* 1998), respectively.

Sometimes fracture fixation is required immediately, leaving no time for expansion of cells in culture. In this case, a method that relies on the ingrowth of bone into the scaffold may be employed. Osteoconduction is the provision of a scaffold to enable bone growth at a time and in a location when and where it would occur because of the presence of bone-forming cells and bioactive molecules from the surrounding tissue. Osteoinduction, on the other hand, is the process of effecting bone growth in a location and at a time where and when it would not occur by providing the scaffold, cells, and/or bioactive molecules (Yaszemski and Yasko 1998). Defects that require osteoinductive materials are called critical sized defects. Rat stromal osteoblasts have been shown to attach to and migrate upon three-dimensional PLGA foams *in vitro* (Ishaug *et al.* 1997) and *in vivo* (Ishaug-Riley *et al.* 1997) in noncritical sized defects, demonstrating the osteoconductivity of this material. PPF has also been shown to be osteoconductive both *in vitro* (Peter *et al.* 2000a) and *in vivo* (Yaszemski *et al.* 1995, Lewandrowski *et al.* 2000). Five weeks after implantation of PPF constructs into noncritical sized rat tibial defects, the degrading scaffolds exhibited ingrowth of new bone trabecuae as well as good integration with the surrounding bone (Yaszemski *et al.* 1995). The addition of peptide sequences such as GRGD (Jo *et al.* 2000) that modulate cellular behavior should augment the ability of PPF to assist in bone regeneration, possibly making it osteoinductive and available for use in critical sized defects.

Another technique employed to enhance bone regeneration in polymeric scaffolds is through the incorporation of bioactive molecules that modulate cellular activity. The most commonly investigated osteogenic factors are the members of the transforming growth factor-β (TGF-β) "superfamily" which includes TGF-β1 and bone morphogenetic protein-2 (BMP-2). One method of enhancing the *in vivo* efficiency of these factors is to facilitate their sustained release over an extended period time. With tissue-engineered devices, there are two different potential delivery systems (Babensee *et al.* 2000). Growth factors can be incorporated directly into the scaffold and released by a diffusion-controlled mechanism regulated by pore morphology as the scaffold degrades. Alternatively, the growth factor delivery system, in the form of microparticles, nanoparticles, or rods can be incorporated into a scaffold. For these devices, the rate of release would depend on the degradation of the delivery device with some contribution from that of the scaffold, and rate of growth factor diffusion through the pores of the delivery system and of the scaffold. For both delivery systems, dosage, release pattern, kinetics of release, and duration of delivery are important parameters that need optimization (Babensee *et al.* 2000).

The use of the poly(α-hydroxy acid) family of polymers has been the main area of focus for growth factor delivery devices. Unilateral critical sized defects in rabbit radii treated with a scaffold of PLA delivering rhBMP-2 demonstrate greater radiopacity as well as greater torque at failure as compared to untreated controls (Wheeler *et al.* 1998). PLGA microparticles loaded with rhBMP-2 incorporated in an allogenic blood clot promote bone formation in rat mandibular defects (Zellin and Linde 1997). Finally, towards development of a fully synthetic carrier/scaffold system, TGF-β1 encapsulated in blended PLGA/poly(ethylene glycol) microparticles was released and promoted proliferation and differentiation of osteoblast precursor cells cultured for 28 days *in vitro* on PPF substrates (Peter *et al.* 2000b).

4. Conclusions

Bone performs many important functions in the body. When bone is deficient in one location, many of these functions can be carried out by bone at distant sites, but its mechanical role in supporting the body and assisting in body movement and locomotion cannot. The mechanical functions, once lost by trauma or disease, can only be regained by restoring skeletal continuity at the location of interest.

Tissue engineering approaches that use polymeric, biodegradable materials hold great promise for regeneration of fully functional bone. Polymeric materials are versatile; their material properties can be tailored to meet specific needs through alterations in polymer composition and structure. Before these techniques can become a reality for patients, however, additional work must be done. Novel biomaterials should be developed that will interact with living tissue and modulate bone formation and repair. Further investigation of the bone growth factors should be done to determine the proper release kinetics most effective for particular applications. Improvements in the carriers for these molecules may be needed to achieve the proper spatial and temporal release to enhance cellular attachment and proliferation. As more information is obtained about the biology of bone fracture and repair, strategies will improve and new approaches will be developed, taking the use of bioresorbable polymers for the regeneration of bone closer to clinical application.

Acknowledgments

This work was supported by the National Institutes of Health (R29-AR42639, R01-AR44381, R01-DE13031, and R01-DE13740). E. L. Hedberg also acknowledges the support of a National Institutes of Health Biotechnology Training Grant (5T32GMO8362-10).

See also: Bone and Natural Composites: Properties*; Bone Mineralization$^{\#}$

Bibliography

Babensee J E, McIntire L V, Mikos A G 2000 Growth factor delivery for tissue engineering. *Pharm. Res.* **17**, 497–504

Behravesh E, Yasko A W, Engel P S, Mikos A G 1999 Synthetic biodegradable polymers for orthopedic applications. *Clin. Orthop. Rel. Res.* **367S**, S118–29

Bostrom R D, Mikos A G 1997 Tissue engineering of bone. Atala A, Moony D J (eds.) *Synthetic Biodegradable Polymer Scaffolds*. Birkhauser, Boston Chap. 12

Bouvier M 1989 The biology and composition of bone. In: Cowin S C (ed.) *Bone Mechanics*. CRC Press, Baca Raton, FL Chap. 1

Breitbart A S, Grande D A, Kessler R, Ryaby J T, Fitzsimmons R J, Grant R T 1998 Tissue engineered bone repair of calvarial defects using cultured periosteal cells. *Plast. Reconstr. Surg.* **101**, 567–74

Buckwalter J A, Glimcher M J, Cooper R R, Recker R 1995 Bone biology: I. Blood supply, cells, matrix, and mineralization. *J. Bone Joint Surg.* **77-A(8)**, 1256–75

Crane G M, Ishaug S L, Mikos A G 1995 Bone tissue engineering. *Nature Med.* **4**, 1322–4

Einhorn T A 1999 Clinically applied models of bone regeneration in tissue engineering research. *Clin. Orthop. Rel. Res.* **367S**, S59–67

He S, Timmer M D, Yaszemski M J, Yasko A W, Engel P S, Mikos A G 2001 Synthesis of biodegradable poly(propylene fumarate) networks with poly(propylene fumarate)-diacrylate macromers as cross-linking agents and characterization of their degradation products. *Polymer* **42**, 1251–60

He S, Yaszemski M J, Yasko A W, Engle P S, Mikos A G 2000 Injectable biodegradable polymer composites based on poly(propylene fumarate) cross-linked with poly(ethylene glycol)-diacrylate. *Biomaterials* **21**, 2389–94

Ishaug S L, Crane G M, Miller M J, Yasko A W, Yaszemski M J, Mikos A G 1997 Bone formation by three-dimensional stromal osteoblast culture in biodegradable polymer scaffolds. *J. Biomed. Mater. Res.* **36**, 17–28

Ishaug-Riley S L, Crane G M, Gurlek A, Miller M J, Yasko A W, Yaszemski M J, Mikos A G 1997 Ectopic bone formation by marrow stromal osteoblast transplantation using poly (D,l-Lactic-*co*-glycolic acid) foams implanted in the rat mesentery. *J. Biomed. Mater. Res.* **36**, 1–8

Jo S, Engel P S, Mikos A G 2000 Synthesis of poly(ethylene glycol)-tethered poly(propylene fumarate) and its modification with GRGD peptide. *Polymer* **41**, 7595–604

Kuhn K D 2000 *Bone Cements*. Springer, Berlin

Lewandrowski K U, Gresser J D, Wise D L, White R L, Trantolo D J 2000 Osteoconductivity of an injectable and bioresorbable poly(propylene glycol-*co*-fumaric acid) bone cement. *Biomaterials* **21**, 293–8

Peter S J, Kim P, Yasko A W, Yaszemski M J, Mikos A G 1999 Cross-linking characteristics of an injectable poly(propylene fumarate)/β-tricalcium phosphate paste and mechanical properties of the cross-linked composite for use as a biodegradable bone cement. *J. Biomed. Mater. Res.* **44**, 314–21

Peter S J, Lu L, Kim D J, Mikos A G 2000aa Marrow stromal osteoblast function on a poly(propylene fumarate)/β-tricalcium phosphate biodegradable orthopaedic composite. *Biomaterials* **21**, 1207–13

Peter S J, Lu L, Kim D J, Stamatas G N, Miller M J, Yaszemski M J, Mikos A G 2000b Effects of transforming growth factor β1 release from biodegradable polymer microparticles on marrow stromal osteoblasts cultured on poly(propylene fumarate) substrates. *J. Biomed. Mater. Res.* **50**, 452–62

Peter S J, Miller S T, Zhu G, Yasko A W, Mikos A G 1998 In vivo degradation of a poly(propylene fumarate)/β-tricalcium phosphate injectable composite scaffold. *J. Biomed. Mater. Res.* **41**, 1–7

Peter S J, Nolley M S, Widmer J E, Merwin M J, Yaszemski M J, Yasko A W, Engel P S, Mikos A G 1997 *In vitro* degradation of a poly(propylene fumarate)/β-tricalcium phosphate composite orthopaedic scaffold. *Tissue Eng.* **3**, 207–15

Temenoff J S, Lu L, Mikos A G 2000 Bone tissue engineering using synthetic biodegradable polymer scaffolds. In: Davis J E (ed.) *Bone Engineering*. University of Toronto, Toronto, Canada in press

Thomson R C, Wake M C, Yaszemski M J, Mikos A G 1995 Biodegradable polymer scaffolds to regenerate organs. *Adv. Polym. Sci.* **122**, 245–74

Vacanti C A, Kim W, Upton J, Vacanti M P, Mooney D, Schloo B, Vacanti J P 1993 Tissue engineered growth of bone and cartilage. *Transplant. Proc.* **25**, 1019–21

Wake M C, Gerecht P D, Lu L, Mikos A G 1998 Effects of biodegradable particles on rat marrow derived stromal osteoblasts *in vitro*. *Biomaterials* **19**, 1255–68

Wheeler D L, Chamberland D L, Schmitt J M, Buck D C, Brekke J H, Hollinger J O, Joh S P, Suh K W 1998 Radiomorphology and biomechanical assessment of recombinant human bone morphogenetic protein 2 and polymer in rabbit radius ostectomy model. *J. Biomed. Mater. Res. (Appl. Biomater.)* **43**, 365–73

Yaszemski M J, Payne R G, Hayes W C, Langer R, Aufdemorte T B, Mikos A G 1995 The ingrowth of new bone tissue and initial mechanical properties of a degrading polymeric composite scaffold. *Tissue Eng.* **1**, 41–52

Yaszemski M J, Payne R G, Hayes W C, Langer R, Mikos A G 1996 Evolution of bone transplantation: molecular, cellular and tissue strategies to engineer human bone. *Biomaterials* **17**, 175–85

Yaszemski M J, Yasko A W 1998 Musculoskeletal tissue engineering for orthopedic surgical applications. In: Patrick Jr. C W, Mikos A G, McIntire L V (eds.) *Frontiers in Tissue Engineering*. Elsevier, New York Chap. II.10

Zellin G, Linde A 1997 Importance of delivery systems for growth-stimulatory factors in combination with osteopromotive membranes. An experimental study using rhBMP-2 in rat mandibular defects. *J. Biomed. Mater. Res.* **35**, 181–90

E. L. Hedberg and A. G. Mikos
Rice University, Houston, Texas, USA

Buildings: Plastics and Composites

The development and application of plastics and reinforced plastic composites as building materials has grown dramatically since their modern introduction in the early part of the twentieth century. This growth has been driven by several key factors inherent in plastics technology including their suitability for mass production, high strength-to-weight ratios, ability to be molded into intricate shapes and patterns, and exceptional resistance to environmental attack. Expanded performance has been achieved by combining plastic resins with various reinforcing materials to form plastic composites. Figure 1 provides a comparison of representative properties of plastics to other types of building materials.

The earliest practical application of plastics in construction was the use of phenolic-based resins in electric insulators in the 1920s. Commercial development of important thermoplastics such as poly(vinyl chloride), polyethylene, and polystyrene took place in the 1930s and 1940s, although their use in building construction was not significant until after World War II. With rapid advances in the scientific understanding of the chemistry of plastics, and an increased recognition of their value in building applications, the plastics industry expanded rapidly over the latter half of the twentieth century. Today, plastics comprise a sizable portion of the building materials market and cover a broad spectrum of applications including architectural uses such as vinyl window parts, weather barriers such as polyvinyl roof membranes, thermal insulation such as polystyrene boards, structural components such as pultruded fiberglass beams and columns, and utility conduit such as polyethylene piping. The sales of plastics resins to the North American building and construction market is estimated at 9 billion tons (8 billion tonnes) for the year 2000 with the growth rate in sales expected to be nearly double the increase in the rate of total construction spending (Chanda and Roy 1987).

The term plastic, referring to a moldable, synthetic material composed of organic molecules forming long molecular chains (i.e., polymers), embraces a multiplicity of compositions, each developed to meet specific engineering requirements and market

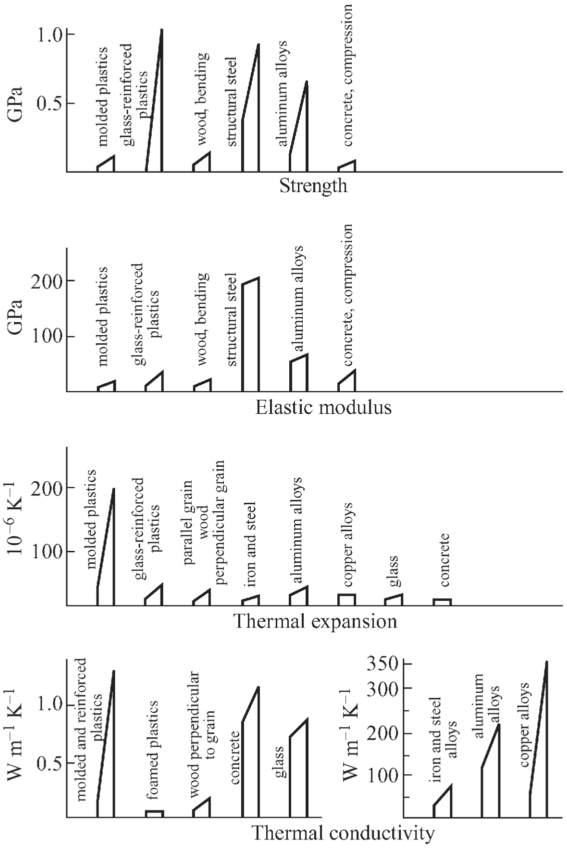


Figure 1
Properties of plastics and other building materials.

demands. Within the two general classes of plastics—thermoplastics and thermosetting plastics—there are dozens of plastic compounds formulated and produced with unique properties for building uses. The versatility is increased with the use of reinforcement materials to form plastic composites. Although sharing a fundamental chemical nature and process, plastic material behavior in response to load and environmental conditions can be as varied as the differences between all other building materials.

The primary advantages of plastics and plastic composites over other building materials are:

- Formability and elimination of joints.
- Light weight and high strength-to-weight ratios.
- Corrosion resistant.
- Nonconductive, nonmagnetic, and insulating.
- Integration of color and texture within the material.
- Tough and wear resistant.

The primary disadvantages of plastics and plastic composites are:

- Loss of properties with high and low temperatures.
- Low stiffness (i.e., high deflection and high creep).
- Combustible and smoke generating.
- Susceptible to static and cyclic fatigue.
- High material and fabrication costs.

1. Properties of Basic Types

1.1 General

Structural plastics are formed from various types of thermoplastic or thermosetting resins. Fibrous reinforcements are sometimes added to the resin substrate to enhance structural properties. This is much more typical with certain thermosetting resins used to form composites with high strength and stiffness. Various additives and fillers may be added to enhance durability, stiffness, fire resistance, or other performance qualities, or to reduce material costs. In view of the large number of available combinations of resin materials, additives, fillers, and reinforcements, typical structural properties given in this section are only intended to illustrate the ranges of potential behavior of some plastic materials that are often found in building applications. Properties are highly influenced by the manufacturing process. Service environments and duration of applied forces also significantly affect properties.

In addition to basic structural properties, selections of specific plastics are related to material cost, formability, availability and cost of processing, directional variations in strength and stiffness, color and surface finish, thermal performance and expansion and contraction, toughness and impact resistance, creep and strength loss with duration of load, unit weight, effects of exposure to service environments (e.g., moisture, chemicals, UV light, etc.), fire resistance, and suitability for property enhancement with additives such as fibrous reinforcements, fillers, UV, or other stabilizers and plasticizers. See Harper (1996), Berins (1991), Seymour (1991), Hancox and Mayer (1994), and the *Modern Plastics Encyclopedia* (1998) for discussions of many of these effects on thermoplastics and thermosetting plastics.

All plastics are plastic at some stage, i.e., they are fluid or soft, and can be cast or shaped in many forms. They are organic materials based upon carbon. They are synthetic as products of chemical processes. Plastics are polymers that are comprised of groups of atoms called monomers, joined together into molecular aggregations. Large groups of such monomers, typical of most plastics, are called high polymers.

1.2 Thermoplastics

A fundamental characteristic of thermoplastics is that they soften repeatedly when heated and harden when cooled. Because of their molecular structure, at high enough temperatures they may melt and at low enough temperatures they may become brittle. These high polymers are monomeric units attached to each other as long-chain molecules. They may be simple unbranched chains or have many branches and intricate entanglements. Chains are attracted to each other by molecular binding forces that are decreased by increasing temperature. Although generally of lower strength and stiffness than thermosetting plastics, thermoplastics are suited to a higher variety of processing techniques and formability, and therefore may provide advantages for economical production of large quantities of complex shapes.

Examples of thermoplastics that are used in the building industry are acrylonitrile–butadiene–styrene (ABS), polyethylene (PE), poly(vinyl chloride) (PVC), acrylics (PMMA), polypropylene (PP), polycarbonate (PC), polyamides (PA) (nylon), fluroplastics (TFE) (Teflon), and polystyrene (PS). Numerous variations of resin formulation, combinations with alloy materials, and additives have been developed and these materials may be manufactured in various forms (i.e., cast, extrusion, fiber, film). Process, additives, alloys, and reinforcements influence structural properties. Some typical structural properties are given in Table 1 to illustrate the general level of structural performance of these example thermoplastics. Hollaway (1994) presents more extensive data for screening and investigating a large number of thermoplastics. However, selection of materials and component design should be based on specific data from the materials supplier.

1.3 Thermosetting Plastics

Thermosetting plastics are distinguished from thermoplastics because they go through a soft plastic stage only once, then harden irreversibly, and cannot again be softened. This is because during curing of the initially soft resins by application of heat, the chain molecules of this class of plastics become crossed-linked to each other by strong chemical bonds or during application of heat and pressure become interlinked in three-dimensional clumps of monomeric units. In some cases water in the form of steam is given off as a by-product of interlinking, a reaction called condensation polymerization.

Examples of thermosetting plastics families that are used in the building industry include unsaturated polyesters, epoxies, and phenolics. Representative strength and stiffness properties for unreinforced casting or molding resins are given in Table 1. However, these plastics are usually used with fibrous reinforcement or fillers. These composites are discussed in Sect. 1.4.

1.4 Reinforced Plastic Composites

The majority of reinforced plastic composites used in building industry applications consist of a polyester resin matrix reinforced with fiberglass. A great range of properties is possible both from resin formulations and types of glass fibers. Three categories of polyesters are common: orthophthalics, isophthalics, and vinyl esters, in ascending order of cost and resistance to weathering, moisture, and chemicals. Curing can be accomplished at ambient temperature or accelerated by the application of heat. Many variations in basic resin formulation, catalysts, accelerators, and additives or fillers are used to enhance special properties, sometimes with undesirable trade-offs in other properties. However, resin manufacturers can engineer polyester systems to produce specific process

Table 1
Representative structural properties.

	Strength (MPa (ksi))		Modulus (MPa (ksi))	
	Tensile	Flexural	Compress	Tensile
Thermoplastics				
PMMA	70 (10)	105 (15)	105 (15)	2800 (400)
PVC	50 (7)	75 (11)	85 (12)	2800 (400)
PA	85 (12)	110 (16)	105 (15)	3000 (440)
PS	40 (6)	60 (9)	85 (12)	3100 (450)
ABS-HI	40 (6)	60 (9)	55 (8)	2100 (300)
PP	35 (5)	50 (7)	50 (7)	1400 (200)
HDPE	30 (4)		20 (3)	1000 (150)
TFE	20 (3)	15 (2)	7 (1)	500 (70)
Thermoset plastics				
Polyester	90 (13)	160 (23)	205 (30)	4100 (600)
Vinyl ester	85 (12)	125 (18)	110 (16)	3500 (500)
Epoxy	70 (10)	115 (17)	140 (20)	3500 (500)
Phenolic	40 (6)	70 (10)	90 (13)	2800 (400)

Table 2
Illustrative structural properties of reinforced thermosetting plastics.[a]

	Fiber content (wt. %)	Tensile strength (MPa (ksi))	Flexural strength (MPa (ksi))	Flexural modulus (MPa (ksi))
Glass-polyester				
Mat (COM)	25	68 (10)	108 (16)	6600 (960)
	30	90 (13)	140 (20)	7200 (1040)
	35	110 (16)	150 (22)	8000 (1160)
Spray-up (CSM)	25	48 (7)	72 (10)	6000 (870)
	30	63 (9)	88 (13)	6900 (1000)
WR + mat (COM)	45	160 (23)	230 (33)	10300 (1500)
UN roving + mat (PUL)	50	210/50	210/70	10300/4800
Longitudinal/transverse		(30/7)	(30/10)	(1500/700)
Carbon-epoxy				
UN (PUL)	70	2000 (290)	1700 (247)	200000 (29000)
Glass-phenolic				
Mat	35	100 (15)	150 (22)	5700 (830)
WR	55	280 (40)	300 (44)	17000 (2500)
UN (PUL)	75		1340 (195)	46000 (6670)

[a] COM = contact molding; CSM = chopped strand mat; WR = woven roving; UN = unidirectional; PUL = pultruded.

and end-performance needs. See the *Modern Plastics Encyclopedia Handbook* (1994) for a more extensive presentation of the use of polyester resins as a matrix resin in fibrous composites. Epoxies, which are generally more costly and difficult to handle than polyesters, are also used as a resin matrix for some reinforced plastic composites, especially those with advanced fiber reinforcements. Compared to polyesters, epoxies may offer higher temperature limits, lower shrinkage, and improved bond and resistance to chemical attack.

Phenolics provide the resin matrix for non- or semistructural laminates that use various fillers to reduce cost and improve performance, or sometimes for fiber-reinforced composites.

The mechanical strength of reinforced plastic composites is primarily dependent on the amount, type, and direction of the fiber reinforcement. The amount is a function of the laminate design and manufacturing process, usually specified as a percentage by volume or by weight. The most common fiber type is E-glass fiber. Special fiber types, found in advanced composites having higher strength or stiffness that are used in a few building applications, are S-2 glass, carbon, and aramid. Fiber reinforcements may be provided in three general arrangements: unidirectional with greatest strength in the direction of the fibers, bidirectional with principal strength in each of two directions of fiber orientation, and multidirectional or isotropic with chopped fibers in random orientations or multiple orientations of continuous fibers in a fabric.

Glass reinforcement can be in the form of continuous rovings and bidirectional unwoven tapes with various proportions of fibers in each direction. Also in frequent use are bidirectional fabrics of woven rovings or cloth, and mats or chopped fibers that are essentially isotropic or with some directional variations in strength produced by bias in orientation. Combinations of the above reinforcements such as alternate layers of mat and woven roving are also used. A thin veil of light fiberglass mat is often used with a gel coat resin as a surfacing layer adjacent to a mold surface.

Reinforced plastics are often used for structural applications that also benefit from their nonstructural qualities. Some typical uses in building are described in Sect. 3. Illustrative strength and stiffness for typical fiber-reinforced plastics (FRPs) and carbon-fiber laminates are given in Table 2.

1.5 Plastic Particulate Composites

Acrylic resins and natural mineral particulates constitute CORIAN and similar commercial, artificial marble-like materials widely used for countertops, sinks, and sculptured castings. Portland cement concrete may be modified with polymers to produce polymer-impregnated concrete, polymer–cement concrete, and polymer concrete.

Polymer-impregnated concrete is produced by drying cured standard concrete, evacuating the pores and impregnating with a monomer subsequently polymerized in place by heat, chemical action, or high-energy radiation. Compressive strengths may be increased fourfold, creep drastically diminished, and resistance to sulfate attack and freezing and thawing greatly enhanced.

Polymer–cement concrete is standard concrete with a monomer or polymer latex added during mixing. It is subsequently handled like standard concrete.

Polymer concrete consists of regular mineral aggregate in a polymer matrix, customarily polyester, instead of Portland cement paste. The ingredients are mixed and cast into forms or molds where the polymer hardens, in hours at room temperature or minutes with heat. Unlike standard concrete, no curing time is needed.

1.6 Laminar Composites

For buildings, high-pressure laminated and structural sandwiches are of most interest. High-pressure laminates are most commonly decorative. They consist of layers of paper impregnated with phenolic resin faced with printed paper (sometimes fabric or wood veneer) impregnated with melamine–formaldehyde and overlaid with a melamine–formaldehyde-impregnated veil, all simultaneously pressed into a decorative laminate, usually subsequently bonded to plywood, particleboard, or other substrates. These laminates are used for counter- and tabletops, door facings, furniture, and cabinetwork. Such laminates are resistant to most household solvents and are harder than common varnishes and lacquers, but not as hard as glass.

Structural sandwiches consist of relatively thin facings of dense, strong, stiff materials bonded to relatively thick, less dense, less strong, and less stiff cores to provide lightweight combinations that are strong and stiff because of their geometry (see Fig. 2). In building panels the combination is also expected to provide thermal insulation, acoustical properties, weather resistance, and resistance to normal wear and tear. Facings most commonly employed include plywood, fiberboard, particleboard, gypsum board, steel, aluminum, high-pressure laminates, and glass-fiber-reinforced plastics. Cores include foamed plastics and honeycombs of paper impregnated with phenolic resin. Adhesives to bond facings and cores are generally based on synthetic resins. Some foamed-in-place plastics, e.g., polyurethanes, are self-adhering. The adhesive must withstand shear and tensile stresses between facing and core.

2. Design Considerations

Plastics and reinforced plastics are unique because they offer the design engineer the opportunity of selecting from a wide variety of resin materials, reinforcements, additives, and manufacturing processes to achieve a desired engineered material that provides specific performance requirements. Furthermore, they offer the following attributes that promote their creative use for structural and semistructural applications in buildings:

Formability. Plastics have no inherent form, so they can often be shaped efficiently for a given application. Form can also help to overcome limitations such as low stiffness (see Fig. 2).

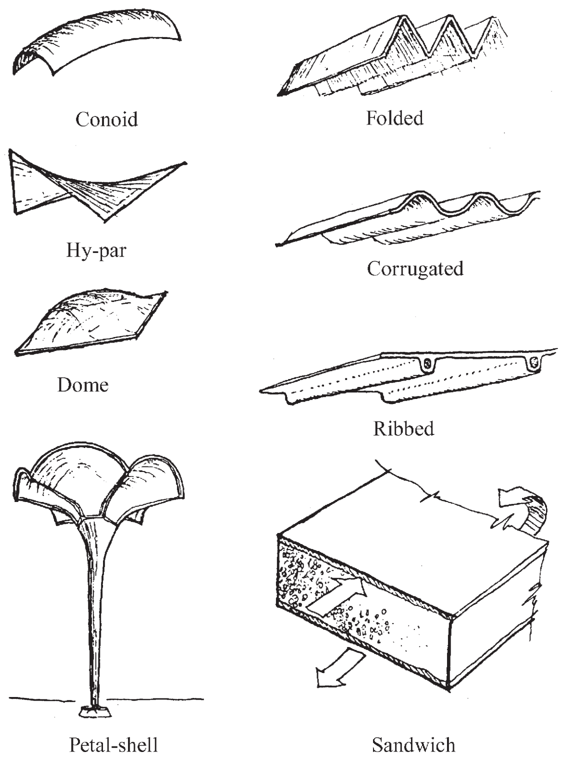


Figure 2
Examples of structural shapes using the formability of plastics.

Consolidation of parts. One-piece forming can often eliminate separate pieces, joints, seams, fasteners, and sealant.

Manufacture or fabrication in thin sections. The manufacturing methods and the toughness of many plastics and composites often allow them to be used in thin sections having efficient structural configurations; by taking advantage of formability, lightweight structural shapes such as shells, folded plates, and tension structures can be achieved. The extremely low specific gravity of plastics foams (0.015 and greater) contributes to low weight in such structures as foam-cored building sandwich panels (see Fig. 2). In sandwiches, facings form a moment-resistant couple; the core provides shear resistance and stabilizes facings.

Light weight and good strength-to-weight ratio. The low specific gravity of most plastics and reinforcing materials results in high ratios of strength to weight.

Corrosion resistance. Because they are largely immune to electrochemical corrosion, plastics and composites can often be used for applications in highly corrosive environments.

Transparency. Certain plastics are highly transparent or translucent to light for use in windows, light fixtures, skylights, and similar applications.

Integral color. Plastics integrally colored with dyes and pigments can be cut and otherwise worked without loss of color. Permanence is largely a matter of the permanence of the pigments and dyes. Surface colors may be applied as paints or gel coats.

Wear resistance and smoothness. Although they may be scratched and indented more easily than glass and other harder materials, plastics often provide good to excellent resistance to wear or abrasion. Nylon hardware parts frequently exhibit outstanding wear resistance and quiet operation. Vinyl flooring is resistant to wear and abrasion. The smooth inner surfaces of extruded pipes result in low resistance to flow of liquids.

Insulation. Many plastics exhibit low thermal conductivity and are used as insulating materials, especially when formed as rigid foams.

Nonconductive. Most plastics are very poor electrical conductors which makes them suitable for applications where electrical insulation and where nonconductive building components are required.

Energy and impact absorption. Fiber reinforcement toughens plastic composites and both rigid and flexible foam plastics absorb impact and blast effects.

Flexibility. Some plastics are inherently soft or flexible; others become so by additives and copolymerization and have both structural and nonstructural uses.

Fire resistance. Plastics are organic materials and can be destroyed by fire. Some burn readily, while others are self-extinguishing when an outside flame is removed. When fire resistance is a consideration, ease of ignition, rate of flame spread, rate of heat release, smoke release, toxicity of products of combustion, amount of combustible material, and structural property loss with temperature rise are among the factors that require consideration. Thermoplastics soften or melt at sufficiently high temperature, whereas thermosets maintain their shapes until, with enough heat, they ignite.

Structural. Unreinforced thermoplastics are used for many structural and semistructural applications in building applications. Generally, their properties are not as great as reinforced composites and feasible manufacturing procedures are suited to production of many similar components of limited size compared to reinforced composites with thermosetting resins. However, a large variety of complex shapes can be produced within the various process limitations. As was emphasized above, the structural properties of fiber-reinforced composites vary with the type, amount, and direction of the reinforcing fibers. Their properties and cost are also closely related to the available methods for fabricating a particular product or component. For example, the amount of fiber and strength of a component fabricated by hand lay-up contact molding may be less than a similar structure fabricated by high-pressure molding, or by pultrusion.

The successful use of plastics and reinforced plastic materials for structural applications in building construction requires a knowledge of performance characteristics in addition to strength and stiffness obtained from standard short-time tests. Unreinforced plastics and the resin matrices of reinforced plastics are basically viscoelastic materials whose strength and stiffness are reduced by creep under various durations of stress. For components subject to loads with significant duration, it is important to know the magnitude of expected creep strain and reduced strength as a function of time under the design maximum stress level (service load stress times a safety factor). The addition of reinforcement generally reduces losses from these viscoelastic effects, but creep and creep rupture still should be determined for reinforced plastics. Since it is expensive and time consuming to obtain adequate engineering data on the effects of long-duration forces on the many varieties of reinforced and unreinforced plastics, such data often must be obtained for design of specific structures subject to long-term loads.

The relatively low stiffness (elastic modulus) and the use of thin shapes with both reinforced and unreinforced plastics require an accurate determination of buckling resistance for elements subject to compression. This determination is more complex for reinforced elements whose elastic modulus often varies with direction. Evaluation and control of deflection also is more significant with these materials.

Thermal coefficients of expansion generally are greater with plastic materials than they are for conventional structural materials. This requires more consideration of the movements produced by thermal change with plastic structures. The thermal coefficient is reduced in the direction of increasing amounts of fiber in reinforced composites. Carbon-fiber composites have very low or negative coefficients of expansion.

Service environments also may greatly affect the structural performance of plastics and reinforced plastics. Strength and stiffness properties generally decrease when subject to increasing temperatures above normal ambient room and outdoor levels and may increase for decreasing temperatures. The effect of temperature on strength and stiffness is a significant criterion when selecting a suitable plastic for a particular application. The effect of outdoor exposure to UV light is another important design consideration. For some materials, UV stabilizers may be added to greatly improve resistance to UV degradation. Also, protective coatings such as gel coats are used. Gel coats and veil fabrics are used to obtain good surface qualities with reinforced plastics.

Applications of plastics and plastic-based composites often take advantage of their superior corrosion resistance in many environments that are corrosive to metals. However, structural properties may be reduced by exposure to moisture or various concentrations of acids or alkalis. Manufacturers and fabricators have a great deal of data on the effects of exposure to various environments and chemicals.

3. Uses in Buildings

3.1 Architectural and Mechanical

These constitute by far the major applications in buildings. In these applications, plastics are not called upon to carry significant loads, and failure does not entail appreciable danger to occupants of buildings. Among the most important of such uses are the following.

(i) *Pipe, fittings, and conduit*. This, a very large class of uses, includes pipe and fittings mainly for drains, vents, and water and underground lines. PVC is most widely used, followed by high- and low-density PE, ABS, reinforced polyester, PP, and PS.

(ii) *Thermal insulation*. Efficient thermal insulation materials are provided by prefoamed or foamed-in-place plastics. Prefoamed plastics, such as PS and polyurethane, are typically blown as large masses and cut into slabs and blocks for use under concrete slabs, in cavity walls, and as sandwich cores. Blocks of foamed-in-place type are also available and provide good insulation if properly formulated and applied.

(iii) *Flooring*. PVC is compounded with fillers, fiber, and additives to provide resilient sheet and tile flooring. Urethane foams are employed as underlayments for rugs. Epoxies provide paving materials.

(iv) *Plumbing*. Bathtubs, shower stalls, and other plumbing fixtures use large quantities of reinforced polyesters in addition to acrylics, polyacetals, thermoplastic polyesters, and PS.

(v) *Glazing and skylights*. Flat, corrugated, and thermoformed acrylic and polycarbonate sheets are widely used, as are flat and corrugated reinforced polyester sheets. Flat reinforced plastic sheets are bonded to metal grids to form translucent structural wall and roof panels.

(vi) *Lighting fixtures*. Acrylics, cellulosics, polycarbonates, PS, and PVC provide a multiplicity of indoor and outdoor lighting fixtures.

(vii) *Moisture barriers*. Low-density PE and PVC find extensive use as barriers to prevent moisture migration through concrete slabs on grade and vapor passage through walls and roofs, and as liners for swimming pools.

(viii) *Profiles*. Window parts, rainwater conduits, moldings, and many others are chiefly extruded PVC and PE, solid or foamed.

(ix) *Wall coverings*. PVC and PS account for most wall coverings, including flexible sheets for walls subject to considerable wear.

(x) *Panels and siding*. These are mainly PVC and reinforced polyesters, with acrylics and cellulose acetate–butyrate also employed.

3.2 Structural

Use of plastics for structural building applications is increasing as more efficient processes for medium- and high-volume production become available. These usually employ plastic-based fibrous composites. Some typical applications are:

(i) Thin-walled structural shapes, such as rectangular and cylindrical tubes, I-shapes, channel and angle shapes, and roof decking manufactured by pultrusion of glass-reinforced polyester are available from stock and special order for primary structure, and wall and roof panels. These are mainly used in buildings and other structures with special requirements for corrosion resistance.

(ii) Sandwich panels with reinforced plastic facings (skins) and foam plastics or plastic-coated paper honeycomb cores. Metal or reinforced concrete are often used for facings of sandwich wall panels with foam plastic insulating cores.

(iii) Shell roof structures of various shapes such as spherical domes, hyperbolic paraboloid, and free-form doubly-curved shapes. Also, sandwich panel shells for larger spans (see Fig. 2).

(iv) Thermoplastics for low-maintenance fencing and small yard buildings.

(v) Thermoformed acrylic (Plexiglas) skylights and polycarbonate (Lexan) window panels.

(vi) Recycled thermoplastic lumber for replacing wood in low-stiffness, rot-resistant applications.

(vii) Plastic-coated fabrics for tensioned roof structures. Vinyl coated nylon is a common example for roofs of air-inflated buildings and pretensioned tents.

3.3 Restoration

Plastic-based composites of various types are widely used in building restoration work. Some typical applications are described below:

(i) Glass-fiber-reinforced polyesters are used for replacing deteriorated stone, terracotta, or concrete ornamental friezes, statues, and other decorative components on building facades.

(ii) Polymer-impregnated concrete is used to replace deteriorated concrete wearing surfaces in parking structures and bridge decks.

(iii) Carbon-fiber/epoxy reinforcement strips are applied to the exterior of deteriorated concrete beams and slabs to restore structural strength.

(iv) Various types of reinforced polyester and epoxy resin reinforcement is wrapped around reinforced concrete to improve ductility and seismic resistance.

3.4 Utility/Corrosion-resistant Structures

Various thermoplastics and glass-fiber-reinforced polyesters are widely used for buried and above-ground structures that require long-term resistance to corrosion. Examples include:

(i) Buried FRP petroleum storage tanks.

(ii) Buried FRP pressure piping and FRP piping that transmit corrosive fluids or operate in corrosive environments such as sewage treatment plants.

(iii) Above-ground FRP tanks and ducts in corrosive environments.

(iv) ABS and PVC drain, vent, and waste piping for domestic and industrial building service.

(v) PVC electrical conduit.

(vi) Gratings, railings, and stairways in place of metal components in sewage treatment plants, industrial plants, and similar facilities with corrosive environments.

3.5 Auxiliaries to Other Materials

Among the principal uses of plastics as auxiliaries to other building materials are: (i) protective and decorative coatings (paints, varnishes, and lacquers); (ii) structural adhesives and resin-bonding agents, especially for wood products; and (iii) building sealants in the form of caulking compounds and gaskets. Contemporary sealants augment the older oil- and resin-based varieties, and may be two-component (mixed on the job) or single-component (ready to use, often as cartridges for sealant guns). Polymers employed include polysulfide rubber, acrylics, silicones, chlorosulfonated PE, and urethanes. Similarly, gaskets draw upon various polymers including natural and synthetic rubbers and PVC.

4. Resources

The designer, specifier, and user of plastics and composite materials will find numerous resources to aid in the selection and application of these materials in building construction. Resin suppliers and plastic product manufacturers are primary sources for test data, design properties, and installation specifications for their products. Governing building codes, reference standards, and third-party design guides provide specific requirements for material selection. Additional resources are available from industry trade associations.

4.1 Codes and Standards

Building codes and design standards for plastics and composites have not advanced to a state comparable to that of more traditional building materials such as wood, steel, and concrete. This is due largely to the multiplicity of plastic formulations and complexity of properties, and to some extent to the proprietary nature of the plastics industry. However, concurrent with the increased use of plastics in building construction, building codes are increasingly incorporating specifications for plastic products and adopting existing national test standards for qualifying plastics materials in building components. This is exemplified in the 1997 Uniform Building Code (UBC) (Uniform Building Code 1997, Vols. 1 and 3). Following are UBC citations for plastic building materials:

Volume 1, Uniform Building Code

- Section 601.5.5. Trim: allowance of foam plastic trim for ceilings and walls.
- Section 1404. Vinyl siding: allowance for vinyl siding on exterior walls complying with UBC Standard 14-2.
- Chapter 26. Plastics: requirements for foam plastics, light-transmitting plastics, and plastic veneers.

Volume 3, Material, Testing and Installation Standards

- UBC Standard 14-2. Vinyl siding.
- UBC Standard 15-6. Modified bitumen, thermoplastic and thermoset membranes use for roof coverings.
- UBC Standard 25-1. Plastic cement.
- UBC Standard 26-3. Room fire test standard for interior of foam plastic systems.
- UBC Standard 26-4. Method of test for the evaluation of flammability characteristics of exterior, nonload-bearing wall panel assemblies using foam plastic insulation.
- UBC Standard 26-5. Chamber method of test for measuring the density of smoke from the burning or decomposition of plastic materials.
- UBC Standard 26-6. Ignition properties of plastics.
- UBC Standard 26-7. Method of test for determining classification of approved light-transmitting plastics.
- UBC Standard 26-8. Room fire test standard for garage doors using foam plastic insulation.

Similar code requirements and standards references can be found in other national building codes, such as BOCA (1996).

Traceski (1990) provides a thorough summary of national and international standards for plastics and composites for all application industries. Some of the major standards-writing organizations and resources for building materials, including plastics, are listed below:

(i) *American Society of Testing and Materials* (*ASTM*). Develops consensus standards for test methods, specifications, recommended practices, nomenclature, and definitions for materials.

Committees:

Committee D20 on plastics

Committee D30 on high modulus fibers and their composites

Committee F17 on plastic piping systems.

Publications:

Annual Book of Standards. Vols. 08.01, 08.02, and 08.03 on plastics and Vol. 08.04 on plastic pipe and building products.

Special Technical Publications (STPs) presenting articles on current research and applications: over 30 volumes on composites and plastics.

(ii) *International Organization for Standards* (*ISO*). An international standards-developing body comprised of national standards bodies from over 90 countries; promotes worldwide standardization.

Committee: ISO Technical Committee 61 on plastics.

Publications:

ISO Standards Handbook 21. Vol. 1, Plastics: terminology, sampling and properties; Vol. 2, Plastics: materials and products.

ISO/TC 61. Plastics: Instruction and Guide Manual, Document No. 570E.

(iii) *German Institute for Standardization* (*DIN*). Most German standards for plastics, textile glass reinforcements, and glass-reinforced plastics have related ISO standards.

(iv) *British Standards Institute*. Promulgates the official standards for the UK. British Standards (BS) for plastics and polymer products are available from the US Department of Commerce, National Institute of Standards and Technology.

4.2 Design Guides

No comprehensive code-sanctioned design standards for plastics and composites exist. However, authoritative design guides have been promulgated by professional organizations and consortiums. Notable examples are:

(i) *Structural Plastics Design Manual* (ASCE 1984).

(ii)The *EUROCOMP Design Code and Handbook* (Clarke 1996).

4.3 Trade Organizations

Trade organizations provide information on manufacturing sources, applications, and design and specification guidelines. Key organizations are listed below:

(i) The Society of the Plastics Industry (SPI), Washington, DC.

(ii) SPI-Composites Institute, Harrison, NY.

(iii) Society of Plastics Engineers (SPE), Brookfield, CT.

(iv) Composites Fabricators Association, Arlington, VA.

Bibliography

ASCE 1984 *Structural Plastics Design Manual*, Manuals and Reports on Engineering Practice no. 63. American Society of Civil Engineers, New York

Berins M L 1991 *Plastics Engineering Handbook*, Society of the Plastics Industry. Van Nostrand Reinhold, New York

Birley A W, Scott M J 1982 *Plastics Materials; Properties and Applications*. Blackie, Glasgow, UK

BOCA 1996 *National Building Code*, 13th edn. Building Officials & Code Administrators International. County Club Hills, IL

Chanda M, Roy S 1987 *Plastics Technology Handbook*. Dekker, New York

Clarke J L (ed.) 1996 Structural design of polymer composites. *EUROCOMP Design Code and Handbook*, European Structural Polymeric Composites Group. E & FN Spon, London

Facts & Figures of the US Plastics Industry 1993 Society of the Plastics Industry

Hancox N, Mayer R 1994 *Design Data for Reinforced Plastics*. Chapman & Hall, London

Harper C A 1996 *Handbook of Plastics, Elastomers and Composites*, 3rd edn. McGraw-Hill, New York

Hollaway L (ed.) *Handbook of Polymer Composites for Engineers*. Woodhead Publishing Ltd., Cambridge, UK

Modern Plastics Encyclopedia '98 1998 McGraw-Hill, New York

Modern Plastics Encyclopedia Handbook 1994 McGraw-Hill, New York

Seymour R 1991 *Reinforced Plastics Properties and Applications*. ASM International, Materials Park, OH

Traceski F T 1990 *Specifications & Standards for Plastics & Composites*. ASM International, Materials Park, OH

Uniform Building Code 1997 International Conference of Building Officials, Whittier, CA

F. J. Heger and P. A. Sharff
Simpson Gumpertz & Heger, Inc., Arlington, Massachusetts, USA

C

Carbon Aerogels

Sol–gel polymerization of metal alkoxides leads to the formation of highly cross-linked and transparent gels. If the gels are dried by hot air, they shrink owing to a large capillary force and the gel structures are often collapsed. To preserve the gel structure and to minimize the shrinkage during drying, it is useful to remove the solvent from the gel under supercritical conditions. The gel dried under these conditions is called an *aerogel.*

Although the most common aerogels are synthesized by sol–gel polymerization of silicon alkoxide, Pekala *et al.* have synthesized organic hydrogels by sol–gel polycondensation of resorcinol (1,3-dihydroxybenzene) with formaldehyde and prepared resorcinol–formaldehyde (RF) aerogels by supercritical drying with carbon dioxide (Pekala 1989, Pekala *et al.* 1992). They have pyrolyzed the aerogels in an inert atmosphere at 1323K obtaining carbon aerogels (Pekala and Alviso 1992). RF and carbon aerogels have a high porosity ($>80\%$), and a high surface area ($400–900\,\mathrm{m^2\,g^{-1}}$). The aerogels are expected to be used as adsorbents, electric double layer capacitors, and materials for chromatographic separation.

The method proposed by Pekala *et al.* is extremely interesting from the viewpoint of developing mesoporous carbons. They have characterized organic and carbon aerogels and reported their electrochemical applications (e.g., Ruben *et al.* 1992, Pekala and Schaefer 1993, Wang *et al.* 1993, Pekala *et al.* 1994, Alviso *et al.* 1997, Pekala *et al.* 1998, Reichenauer *et al.* 1998). Tamon *et al.* have experimentally elucidated the influence of the amounts of resorcinol, water, and basic catalyst used in the sol–gel polycondensation on the porous structures of RF and carbon aerogels (Tamon *et al.* 1997, Tamon and Ishizaka 1998a). They have also proposed a control procedure of porous structures of organic and carbon aerogels (Tamon *et al.* 1998). They have also confirmed that carbon cryogels prepared via sol–gel polycondensation, freeze drying, and pyrolysis have a similar porous structure to carbon aerogels (Tamon *et al.* 1999).

1. *Synthesis of Carbon Aerogels*

The synthesis procedure of RF hydrogels is shown in Figure 1. When resorcinol (R) reacts with formaldehyde (F), RF hydrogels are synthesized via polycondensation, cluster formation, and gelation (Pekala 1989). Here, sodium carbonate (C) and ionized and distilled water (W) are used as a basic catalyst and a diluent, respectively. RF aerogels are prepared by supercritical drying with carbon dioxide and carbon aerogels are obtained by pyrolyzing the aerogels.

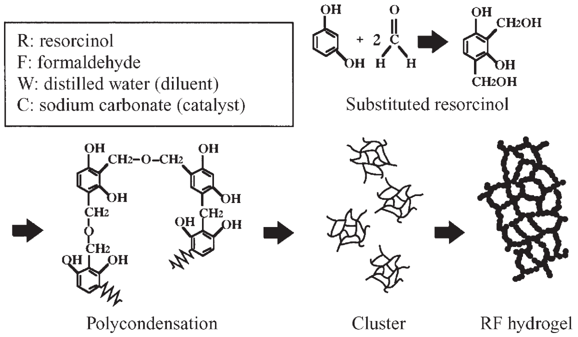


Figure 1
Synthesis of resorcinol–formaldehyde hydrogel.

The small-angle x-ray scattering (SAXS) technique has been applied to the gelation process of RF hydrogels (Tamon *et al.* 1998b). The structure formation of the hydrogels during the sol–gel transition is revealed by applying Guinier and power-law equations to the SAXS data. At the initial stage of the synthesis of RF hydrogels, small clusters of approximately 2 nm consisting of branched polymeric species form that show a mass fractal dimension. The clusters aggregate and form particles of about 3–6 nm that show a surface fractal dimension. The hydrogel structure is fixed by gelation and the particles grow to approximately 4–7 nm. Finally the surface becomes smooth by aging. The pores in RF aerogels are the voids formed by the interconnected particles because the surface area of the particles roughly coincides with the BET (Brunauer–Emmett–Teller) surface area of aerogel determined by applying the BET equation to the adsorption isotherm of nitrogen at 77K.

2. *Porosity of Carbon Aerogels*

Table 1 shows the BET surface area S_{BET}, the mesopore volume V_{mes}, the peak radius of pore size distribution r_{peak}, the surface area of micropores S_{mic}, the micropore volume V_{mic}, and the macropore volume V_{mac} of carbon aerogels. The mole ratio of resorcinol to basic catalyst R/C, the ratio of resorcinol to water R/W, and the mole ratio of resorcinol to formaldehyde R/F used in the synthesis of RF hydrogels are shown in this table. One can see that S_{BET} and V_{mes} are much larger than S_{mic} and V_{mic} of the

---

* Cross references marked by an asterisk are included in this volume.

# Cross references marked by a hash can be found by consulting the Encyclopedia of Materials: Science and Technology.

Table 1
Porous properties of carbon aerogels and cryogels.

Sample	R/C	R/W (gcm^{-3})	R/F	Gel	S_{BET} (m^2 g^{-1})	V_{mes} (cm^3 g^{-1})	r_{peak} (nm)	S_{mic} (m^2 g^{-1})	V_{mic} (cm^3 g^{-1})	V_{mac} (cm^3 g^{-1})
A1	25	0.125	0.5	aerogel	774	1.34	3.6	138	0.04	0.18
				cryogel	1107	1.54	6.2	631	0.22	0.78
A2	25	0.25	0.5	aerogel	749	0.89	2.5	203	0.06	ND
				cryogel	881	0.55	2.5	648	0.22	0.26
B1	200	0.125	0.5	aerogel	575	1.45	9.6	126	0.04	2.79
				cryogel	1203	1.40	5.5	822	0.27	2.61
B2	200	0.25	0.5	aerogel	1215	2.44	6.2	820	0.26	ND
				cryogel	908	1.39	5.5	549	0.18	0.44

R/C: mole ratio of resorcinol to basic catalyst; R/W: ratio of resorcinol to water; R/F: mole ratio of resorcinol to formaldehyde; ND: not detected.

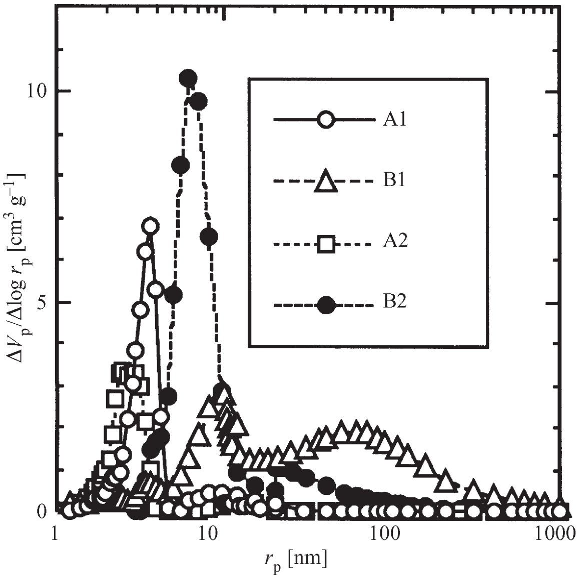


Figure 2
Pore size distribution of carbon aerogels (reproduced by permission of Elsevier Science from *Carbon*, 1998, 36, 1398).

aerogels except for B2. It is also interesting that the V_{mes} of all the aerogels except B1 are larger than V_{mac}. Consequently, it is considered that the carbon aerogels prepared under the conditions of R/W$\geqslant$0.25 have few macropores. This is because the mesopores and macropores are the gel voids formed by the interconnected particles and the voids depend on the amount of distilled water used as diluent in the synthesis of the RF hydrogels.

Figure 2 shows pore size distributions of the carbon aerogels determined using nitrogen adsorption (radius $r_p \cdot$ ca. 25 nm) and mercury porosimetry (radius $r_p \cdot$ ca. 10 nm) methods (Tamon and Ishizaka 1998a). One can see that the aerogels except B1 have sharp mesopore size distributions. On the other hand, B1 has a broad distribution of $r_p = 5$–400 nm. For the same R/W, the peak radius of the distribution r_{peak} increases with increasing R/C. For the same R/C, r_{peak} increases with decreasing R/W.

3. Control of Porous Structure

Mesoporous structures of carbon aerogels can be controlled by choosing synthesis conditions of RF hydrogels (Tamon *et al.* 1998). First, low R/C should be chosen to prepare RF aerogels with sharp mesopore size distributions. For example, RF aerogels of $r_{peak} = 2.5$–4.0 nm are obtained by adjusting R/W at R/C = 12.5. To prepare RF aerogels of larger pore radius, R/W is changed at larger R/C. One can prepare the aerogels with r_{peak} ranging from 2.9 nm to 6.0 nm at R/C = 25, from 3.9 nm to 6.9 nm at R/C = 75, and from 4.5 nm to 9.2 nm at R/C = 100. Hence, RF aerogels of $r_{peak} = 2.5$–9.2 nm can be prepared by this procedure. To prepare the aerogels of $r_{peak} < 2.5$ nm, the amount of distilled water used in the sol–gel polycondensation should be decreased but resorcinol cannot be dissolved in the solution. When trying to prepare RF aerogels of $r_{peak} > 9.2$ nm, R/C must be > 100. In this case, other peaks appear in pore size distributions (Tamon *et al.* 1997).

Although RF aerogels shrink by 1–4 nm during pyrolysis, the shape of the pore size distribution is retained and r_{peak} can be controlled in the range of 2.0–6.1 nm by changing R/C and R/W. The following equations are useful to control r_{peak} and V_{mes} of carbon aerogels:

$$\ln r_{peak} = 0.226\ln(\mathrm{R/C}) - 0.452\ln(\mathrm{R/W}) - 0.254$$

$$(\mathrm{R/F} = 0.5, 12.5 \cdot \mathrm{R/C} \cdot 100, 0.125 \cdot \mathrm{R/W}0.5, \cdot = 0.953) \quad (1)$$

$$\ln V_{\mathrm{mes}} = 0.406\ln(\mathrm{R/C}) - 0.812\ln(\mathrm{R/W}) - 2.476$$
$$(\mathrm{R/F} = 0.5, 12.5 \cdot \mathrm{R/C} \cdot 75, 0.125 \cdot \mathrm{R/W} \cdot 0.5, = 0.974) \quad (2)$$

where · is a correlation coefficient.

4. *Carbon Cryogels*

Although carbon aerogels are unique materials with high surface areas and large mesopore volumes, the cost of supercritical drying is extremely high. Hence, RF cryogels are prepared by freeze drying with *t*-butanol and carbon cryogels are obtained by pyrolyzing the cryogels (Tamon *et al.* 1999). Table 1 also shows porous properties of carbon cryogels. The values of V_{mes} of the cryogels are around 60–120% of those of carbon aerogels. Carbon cryogels are mesoporous materials with high surface areas $>800\,\mathrm{m}^2\,\mathrm{g}^{-1}$ and large mesopore volumes $>0.55\,\mathrm{cm}^3\,\mathrm{g}^{-1}$.

5. *Conclusion*

Carbon aerogels are mesoporous materials with high surface areas and large mesopore volumes. The mesoporous structure of the aerogels can be controlled by choosing the synthesis conditions of RF hydrogels. Carbon cryogels also have similar porous properties to those of carbon aerogels. Various practical applications of carbon gels should be studied in the future.

Bibliography

Alviso C T, Pekala R W, Gross J, Lu X 1996 Resorcinol–formaldehyde and carbon aerogel microspheres. *Mater. Res. Soc. Symp. Proc.* **431**, 521–6

Pekala R W 1989 Organic aerogels from the polycondensation of resorcinol with formaldehyde. *J. Mater. Sci.* **24**, 3221–7

Pekala R W, Alviso C T 1992 Carbon aerogels and xerogels. *Mater. Res. Soc. Symp. Proc.* **270**, 3–14

Pekala R W, Alviso C T, Kong F M, Hulsey S S 1992 Aerogels derived from multifunctional organic monomers. *J. Non-Cryst. Solids* **145**, 90–8

Pekala R W, Farmer J C, Alviso C T, Tran T D, Mayer S T, Miller J M 1998 Carbon aerogels for electrochemical applications. *J. Non-Cryst. Solids* **225**, 74–80

Pekala R W, Mayer S T, Kaschmitter J L, Kong F M 1994 *Sol–Gel Processing and Applications*. Plenum, New York pp. 368–77

Pekala R W, Schaefer D W 1993 Structure of organic aerogels. 1. Morphology and scaling. *Macromolecules* **26**, 5487–93

Reichenaeur G, Emmerling A, Fricke J, Pekala R W 1998 Microporosity in carbon aerogels. *J. Non-Cryst. Solids* **225**, 210–4

Ruben G C, Pekala R W, Tillotson T M, Hrubesh L W 1992 Imaging aerogels at the molecular level. *J. Mater. Sci.* **27**, 4341–9

Tamon H, Ishizaka H 1998a Porous characterization of carbon aerogels. *Carbon* **36**, 1397–9

Tamon H, Ishizaka H 1998b SAXS study on gelation process in preparation of resorcinol–formaldehyde aerogel. *J. Colloid Interface Sci.* **206**, 577–82

Tamon H, Ishizaka H, Araki T, Okazaki M 1998 Control of mesoporous structure of organic and carbon aerogels. *Carbon* **36**, 1257–62

Tamon H, Ishizaka H, Mikami M, Okazaki M 1997 Porous structure of organic and carbon aerogels synthesized by sol–gel polycondensation of resorcinol with formaldehyde. *Carbon* **35**, 791–6

Tamon H, Ishizaka H, Yamamoto T, Suzuki T 1999 Preparation of mesoporous carbon by freeze drying. *Carbon* 37, 2049–55

Wang J, Angnes L, Tobias H, Roesner R A, Hong K C, Glass R S, Kong F M, Pekala R W 1993 Carbon aerogel composite electrodes. *Anal. Chem.* **65**, 2300–3

H. Tamon
Department of Chemical Engineering, Kyoto University, Japan

Carbon Blacks

Carbon black is produced by the incomplete combustion of hydrocarbons such as those found in petroleum oil or natural gas. In its basic form, the smallest primary units of carbon black are aggregates, which in turn comprise partially "coalesced" colloidal sized spheroidal particles with a paracrystalline graphitic microstructure (Fig. 1).

1. *Historical Perspective*

For many thousands years carbon black was used exclusively as a pigment. However, in the early part

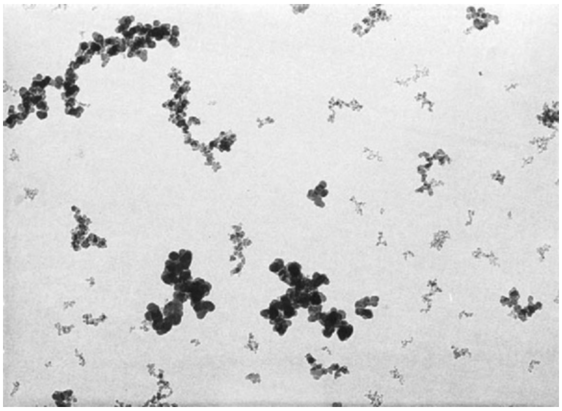


Figure 1
Transmission electron micrograph of carbon black (× 100 000). Note that carbon black is a distributional material.

of the twentieth century its utility broadened, and it is now used in a wide variety of applications. Some examples of the applications of carbon black today include use as a reinforcing agent in tires and rubber goods, as a pigment in inks, paints, and toners, an electrical conductor, and for ultraviolet light protection in plastics.

1.1 The Worldwide Market

In 1997 the worldwide demand for carbon black was 6.36 million tonnes (Columbian Chemicals Company estimate). This demand is expected to grow by over 15% by 2002. Of this volume, most carbon black, 93%, is used in rubber-related applications (tires—71%, mechanical rubber goods—22%). The remaining 7% is used in nonrubber applications such as in plastics, inks, and paints. Globally, the usage of carbon black is split as follows: Americas—35%, Asia—35%, Europe/Africa—30%, with demand in the Asian region expected to grow rapidly during the initial years of the twenty-first century.

2. Production Methods

A number of other processes have been used or are still used today, including (Kuhner and Voll 1993):

(i) The thermal black process—utilizing natural gas and oil as the feedstock. This produces very large primary particles of the order of ~300 nm mean diameter (other processes: ~10–100 nm).

(ii) The acetylene black process—utilizing acetylene gas as the feedstock. Excess heat produced during the formation process results in a less active surface. Carbon black produced by this method is primarily used in conductive applications.

(iii) Channel black process—utilizing natural gas as the feedstock. This process produces a highly oxidized surface, and sees limited use today.

(iv) The lamp black process—utilizing a petroleum oil feedstock. This process is essentially obsolete.

(v) The furnace black process (Fig. 2). Since its introduction in the 1940s, the furnace process has become the predominant method of carbon black production. The furnace itself is a heated, cylindrical, refractory-lined vessel into which a feedstock oil (such as petroleum, coal tar, or ethylene tar) is injected. Process conditions such as temperature, air-to-oil ratio and oil spray position are controlled to allow for partial combustion of the oil, and to manipulate the fundamental properties of the carbon black. After the furnace, this process is not significantly different from other carbon black processes, with heat exchangers (to remove heat from the process stream and aid in heating the furnace), bag collectors (to separate the carbon black from the process/tail gases and begin densification of the carbon black smoke stream), pelletizing/beading

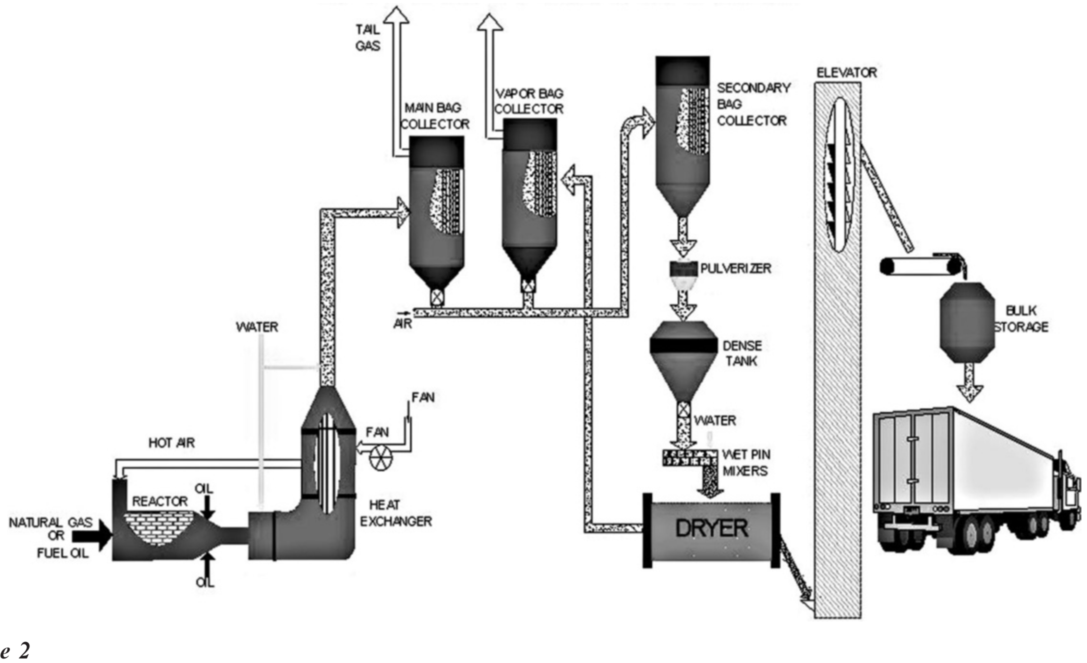


Figure 2
The oil furnace carbon black process.

equipment, dryers, and various packaging, loading, and storage equipment.

3. Carbon Black Formation

Though not completely understood, carbon black formation is believed to occur through the following phases:

(i) atomization and vaporization of the feedstock oil (Fig. 3(a));

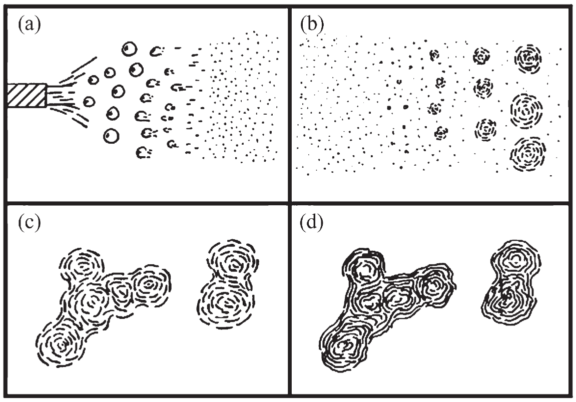


Figure 3
The four primary phases of carbon black formation.

(ii) nucleation of the primary particles and initial surface growth (Fig. 3(b));

(iii) collision of primary particles to form aggregates, with continued surface growth (Fig. 3(c));

(iv) dehydrogenation, carbonization, and possible secondary oxidation (leading to porosity) until the reaction is quenched (Fig. 3(d)).

4. Fundamental Properties of Carbon Black

4.1 Particle Size and Distribution (Fineness)

Particle size can be measured accurately only by electron microscopy (e.g. according to ASTM D3849) (Fig. 4). However, other tests such as tint, iodine number, and nitrogen surface area can also give indications of relative particle size. Typical mean particle size ranges for furnace process grades are 10–100 nm.

Particle size distribution can also play a role in carbon black performance. For example, for a certain mean particle size, a broader particle size distribution will give a lower surface area and tint, resulting in lower viscosity and jetness in plastics and coatings, and lower hysteresis properties in rubbers.

4.2 Aggregate Size and Distribution (Structure)

Aggregate size and distribution properties are also best measured by electron microcopy (ASTM

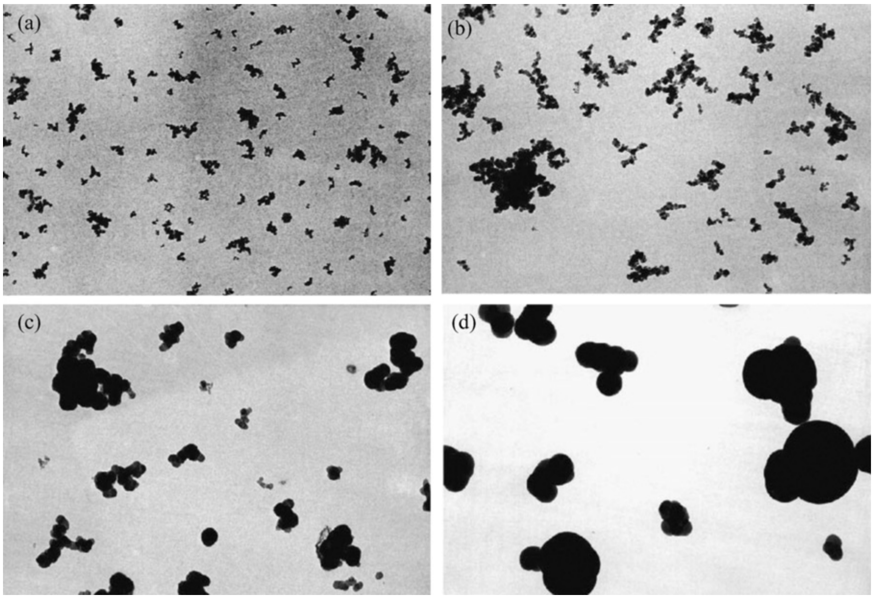


Figure 4
Transmission electron micrographs depicting various grades of carbon black (× 20 000; mean particle sizes in parentheses): (a) N110 (18 nm), (b) N351 (35 nm), (c) N762 (100 nm), and (d) N990 (300 nm).

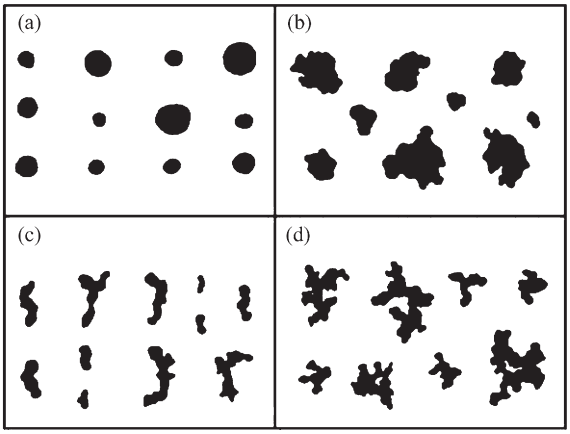


Figure 5
Four classes of carbon black aggregate morphology.

D3849). However, dibutylphthalate absorption (DBPA) (ASTM D2414) is also commonly used to determine the structure of carbon blacks relative to one another.

Aggregates can be categorized into four basic types: spheroidal, ellipsoidal, linear, and branched (Hess and Herd 1993) (Fig. 5). Every carbon black contains a distribution of each of these types, with the higher structure grades containing more branched and linear aggregates, and the lower structure grades containing more ellipsoidal and spheroidal aggregates.

It is also interesting to note that the linear and branched aggregates tend to break down to some degree during intense mixing. However, no aggregate can be broken down to individual primary particles.

Research has shown that aggregates exhibit an anisotropic morphology, tending to be somewhat more two-dimensional or planar in nature than originally believed. Better understanding of these characteristics may give better insight on carbon black formation and performance.

4.3 Porosity and Pore Size Distribution

Most porosity in carbon blacks is characterized as microporosity, or pores less than 2 nm in diameter. Techniques for measuring porosity in carbon black are based on gas absorption or density. Information derived from nitrogen surface area measurement data is one common way to estimate the relative porosity of a carbon black (ASTM D5816).

4.4 Surface Activity and Activity Distribution

Surface activity is a measure of the interaction between carbon black and the system into which it is mixed. It is an exceedingly complex phenomenon in strongly interactive systems such as common elastomers. For weaker interactions, such as in inks and plastics, it may be related to surface oxides. Manufacturers may intentionally oxidize the surface of carbon black particles for these special applications.

5. *Common Characteristics of Carbon Black—Definition, Testing, and Relation to Fundamental Properties*

5.1 Tint (ASTM D3265)

Tint strength is mainly a function of particle size, with smaller particles generally giving a higher tint. At constant particle size, DBPA can influence tint (i.e. lower DBPA gives higher tint). Higher tint carbon blacks generally give higher jetness and viscosity.

5.2 Iodine Number (ASTM D1510)

Iodine number generally increases with increasing surface area, although residual oil on the surface can suppress the iodine number. Higher iodine number carbon blacks generally give higher tinting strength.

5.3 Nitrogen Surface Area (ASTM D4820)

Nitrogen surface area generally increases with decreasing particle size and increasing porosity. Higher surface area carbon blacks generally give higher tinting strength.

5.4 DBPA (ASTM D2414)

DBPA provides an indication of the average aggregate size and aggregate void volume of the carbon black. Higher DBPA indicates larger aggregate size, and generally gives increased viscosity and dispersion, and lower tinting strength.

5.5 Volatile Level (ASTM D1509)

Volatile level gives an indication of functional groups on the carbon blacks surface.

5.6 Impurities

Impurities in carbon black are commonly measured by ASTM D1514 (water wash residue) or by ASTM D1506 (ash content). Common impurities in carbon black include:

(i) refractory—from the brick furnace lining;
(ii) metal—from other process equipment;
(iii) coke—hard carbonaceous particles typically resulting from poorly atomized feedstock oil;
(iv) ash—residual inorganic material from the feedstock oil or water used in the process (for quench, beading, etc.).

6. Typical Applications of Carbon Black

6.1 Rubber

As mentioned above, ~93% of carbon black usage is in rubber applications, and can be classified into two main categories: tires and mechanical rubber goods (e.g. automotive belts and hoses). The reinforcing properties of carbon black provide enhanced performance and durability to the rubber compounds in which they are used. Carbon blacks used in rubber applications are typically categorized as N100–N900 series blacks, where increasing N number correlates with decreasing nitrogen surface area (ASTM D1765).

The influence of carbon black fundamental properties on in-rubber performance is as follows:

- Smaller particle size:
  - increases tensile strength;
  - increases abrasion resistance;
  - reduces rebound;
  - reduces dispersibility.
- Higher DBPA or structure:
  - increases modulus;
  - increases hardness;
  - increases viscosity;
  - reduces die swell;
  - increases dispersibility.
- Higher porosity:
  - increases tear strength;
  - increases conductivity;
  - reduces rebound.
- Higher surface activity:
  - increases abrasion resistance;
  - increases rebound;
  - increases modulus.

6.2 Nonrubber (Plastics, Inks, Coatings, etc.)

The remaining ~7% of carbon usage falls into a broad group of applications, of which the most predominant are plastics, inks, and coatings. Additional applications are varied, and include sealants, toners, and batteries. In these applications, carbon black is used for such properties as appearance (blackness, tint, gloss), conductivity, and ultraviolet light protection.

The influence of carbon black fundamental properties on nonrubber applications is generally as follows:

- Smaller particle size:
  - increases blackness and tinting strength;
  - increases viscosity;
  - increases ultraviolet protection;
  - increases conductivity;
  - decreases dispersibility.
- Higher DBPA or structure:
  - increases viscosity and vehicle demand;
  - increases dispersibility;
  - increases conductivity.
- Higher porosity:
  - increases viscosity and vehicle demand;
  - allows reduced loadings in conductive applications.
- Higher surface activity:
  - improves wetting, microdispersion, and stability;
  - reduces viscosity of liquid systems.

Bibliography

Donnet J-B, Voet A 1976 *Carbon Black*. Dekker, New York
Gruber T C, Herd C R 1993 Rubber Chem. Technol. **70**, 727–46
Hess W M, Herd C R 1993 Microstructure, morphology and general physical properties. In: Donnet J-B, Bansal R C, Wang M-J (eds.) *Carbon Black*, 2nd edn. Dekker, New York
Kuhner G, Voll M 1993 Manufacture of carbon black. In: Donnet J-B, Bansal R C, Wang M-J (eds.) *Carbon Black*, 2nd edn. Dekker, New York

E. B. Sebok and R. L. Taylor
Columbian Chemicals Company, Marietta, Georgia, USA

Carbon Fibers

Carbon fibers have proved to be the main reinforcement for advanced composites for a wide range of applications since the 1960s, but the majority of products still belong to high technology space and aeronautics (Fitzer 1985, Donnet and Bansal 1984). Carbon fibers are a few microns thick, light-weight, very strong, and stiff black synthetic fibers with long aromatic molecular chains, comprising mainly carbon. These fibers are capable of maintaining this structure under extreme conditions of temperature and pressure and, therefore, can be used with all types of matrices, i.e., polymers, ceramics, and metals employing different composite processing techniques.

Though the historical development of carbon fibers goes back to 1879 when produced by Edison for the electric light bulb, the interest in carbon fibers as a reinforcement for composites for structural applications started in the 1960s with the demand from the aeronautical sector for light-weight, strong, and stiff materials. The elastic constant of the graphite lattice is 1060 GN m^{-2} along the layer planes. In order to realize this property in bulk graphite, the graphite material should be in an ultra-thin form, i.e., fibers with layer planes oriented along their length. This generated interest in the development of high-performance carbon fibers with a perfect graphitic structure with the layer planes oriented along the fiber axis.

1. Carbon Fiber Processing

Carbon fibers are produced by the controlled pyrolysis of a carbonaceous material in fibrous form. Therefore, though basically any fibrous material made from a carbonaceous precursor can be used to produce carbon fibers, only rayon, polyacrylonitrile (PAN), and pitch-based fibrous precursors have been used for producing commercial carbon fibers (Fitzer and Heine 1988). A schematic comparison of the processing routes for carbon fibers from these three precursors is shown in Fig. 1. Precursor weight loss and retention of orientation of the graphitic planes had been the central issue in the processing and economics of carbon fiber technology. Recently the carbonization behavior of many other polymeric materials has been studied as potential precursors for carbon fibers. These include phenolics, polyimides, polyamides, polyphenylene, polybenzimidazole, polyethylene, polypropylene, and polyvinylchloride (Bahl *et al.* 1998). Though some of these polymers exhibit high char yield, the properties and structure of the ultimate carbon fibers is not suitable for their application as a high-performance material. However, phenolic-based carbon fibers are used to produce activated carbon fibers.

1.1 Rayon-based Carbon Fibers

Rayon-based carbon fibers were the first to be commercialized and remained dominant up to the early 1970s. Though at present the majority of carbon fibers produced worldwide are derived from PAN and

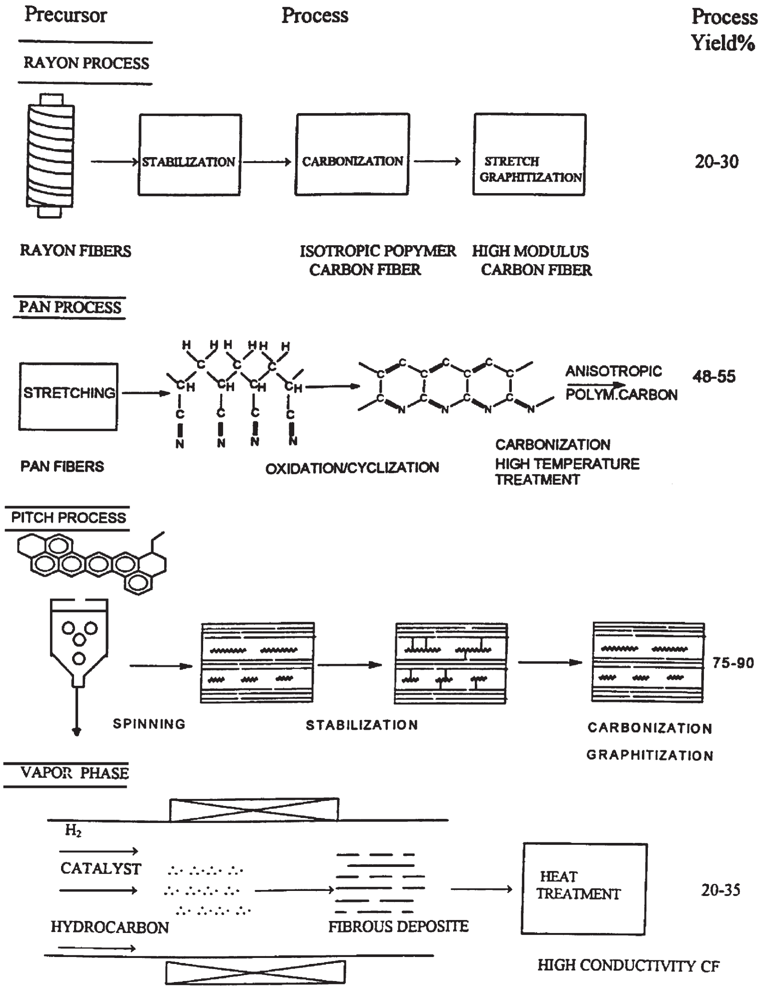

Precursor
Process
Process Yield%
RAYON PROCESS
STABILIZATION
CARBONIZATION
STRETCH GRAPHITIZATION
20-30
RAYON FIBERS
ISOTROPIC POPYMER CARBON FIBER
HIGH MODULUS CARBON FIBER
PAN PROCESS
STRETCHING
ANISOTROPIC POLYM.CARBON
48-55
PAN FIBERS
OXIDATION/CYCLIZATION
CARBONIZATION HIGH TEMPERATURE TREATMENT
PITCH PROCESS
SPINNING
STABILIZATION
CARBONIZATION GRAPHITIZATION
75-90
VAPOR PHASE
H₂
CATALYST
HYDROCARBON
FIBROUS DEPOSITE
HEAT TREATMENT
20-35
HIGH CONDUCTIVITY CF

Figure 1
Schematic comparison of processing routes for carbon fibers.

pitch precursors, the pyrolysis of cellulosic fibers and the development of high-modulus carbon fibers from them is quite significant from the point of view of basic carbonization chemistry. A special grade of rayon fiber with a high degree of crystallinity containing no defects or inclusions is used for carbon fiber manufacture. The processing of carbon fibers from rayon consists of a low temperature (200–400 °C) heat treatment, preferably in air to stabilize the fibers, followed by pyrolysis to 1000 °C in an inert atmosphere (Bacon 1973). During the early stages of decomposition of levoglucosan there is an evolution of H_2O, CO, CO_2, CH_2, etc., and in the later stage a rearrangement of the planes occurs. These are the most important steps in the production of carbon fibers from a rayon precursor. The highly oriented crystalline structure of the precursor fiber is lost due to cleavage of glycosidic linkages during pyrolysis, resulting in a linear shrinkage of 20–25% and a cross-sectional shrinkage of 40–60%. Therefore, carbon fibers (1000 °C) from viscose rayon are low density (1.6–1.7 $\mathrm{g\,cm^{-3}}$) with a very low strength of 400–700 MPa and a modulus of 30–50 GPa. These fibers are good enough for nonstructural applications only, such as ablative applications or as activated carbon fibers. Treatment of the precursors with Lewis acids or ammonia inhibits excessive degradation of levoglucosan, lowers the tar formation, and increases the char yield.

Properties of ex-rayon carbon fibers are improved on further heat treatment to 2500–3000°C especially if the heat treatment is carried out under tension (Bacon and Schalamon 1969). The fibers exhibit plasticity in this temperature range and, hence, orientation of the graphitic layers can be achieved through application of stress, i.e., stress graphitization. Although these carbon fibers posses good mechanical properties ($\sigma = 3.0$–3.5 GPa and $E = 600$–650 GPa) they are very expensive.

1.2 Carbon Fibers from Acrylic Precursors

(*a*) *Polyacrylonitrile precursor*. Among all possible carbon fiber precursors, PAN turned out to be the most suitable precursor from the point of view of overall carbon content (67%), high carbon yield ($\sim$54%, more than 80% on the basis of carbon content), processing, structure, properties, and cost. As a result, PAN-based carbon fibers account for more than 70% of the world's total carbon fiber production and are used in the majority of current advanced composites.

The first extensive study on carbonization of PAN fibers was reported by Shindo in Japan in 1961. However, the breakthrough in the development of high quality carbon fibers came from a British team. Johnson *et al.* (1966), through their experiments on the stabilization of PAN under stress, demonstrated that molecular arrangement and, hence, carbon fibers with high strength and modulus, are obtained through stretching the fibers during stabilization.

PAN is an atactic, linear polymer containing highly polar nitrile pendant groups. The PAN fibers suitable for conversion into high-strength or high-modulus carbon fibers are high molecular weight copolymers having an average molecular weight of $\sim 2.6 \times 10^5$ containing more than 90% acrylonitrile. The comonomers are usually methyl acrylate, methacrylic acid, acrylamide, quaternary ammonium salts of amines, or ethyl methyl propenoate with some itaconic acid or sodium styrene sulfonate (Fitzer 1989a, Edie and Diefendorf 1993). However, the last of these is found to limit the properties of the carbon fiber products at high-temperature treatment. Therefore, PAN precursors with about 2–3% itaconic acid are preferred. The compositions of the PAN precursor, as well as the spinning solution and other spinning conditions, have a great influence on the final carbon fiber properties.

The fundamental fiber structure needed to develop high strength and stiff carbon fibers is created during the initial fiber forming step. Wet, dry, and melt spinning processes have been used to spin PAN precursor fibers. PAN fibers spun from inorganic solvents are found to be, in many respects, superior precursors for the production of carbon fibers. Wet spinning and dry spinning result in circular and dog bone shaped fibers whereas multilobal cross-section fibers can be formed by melt spinning. In wet spun fibers, PAN molecules are organized into fibrils which are generally oriented parallel to the fiber axis and are joined together in a 3D network. This fibrillar network appears to be the precursor for the graphene network that develops during final heat treatment. Leading carbon fiber manufacturers now use dry-jet-spun PAN fibers which result in carbon fibers with much higher mechanical properties than those from wet-spun fibers. Moreover, to produce high quality carbon fibers, the starting PAN precursor fiber should have small diameter, maximum possible orientation of the molecular chains along the fiber axis, maximum crystalline content, and low activation energy for cyclization. Improvement in the precursor fiber can be achieved by post-spin stretching of PAN fibers before stabilization. PAN fibers may be stretched up to 200% during spinning and this improves the structure of the fibers, especially the orientation of molecular chains with respect to the fiber axis, ultimately leading to carbon fibers with better mechanical properties. Therefore, for making carbon fibers, PAN precursors are made under specific controlled conditions and are termed special acrylic fibers (SAF). The precursor cost adds to the end cost of the carbon fibers. Though each producer has its own propriety technology, the basic process involves three steps; preoxidation stabilization, carbonization, and a high-temperature heat treatment. Figure 2 shows

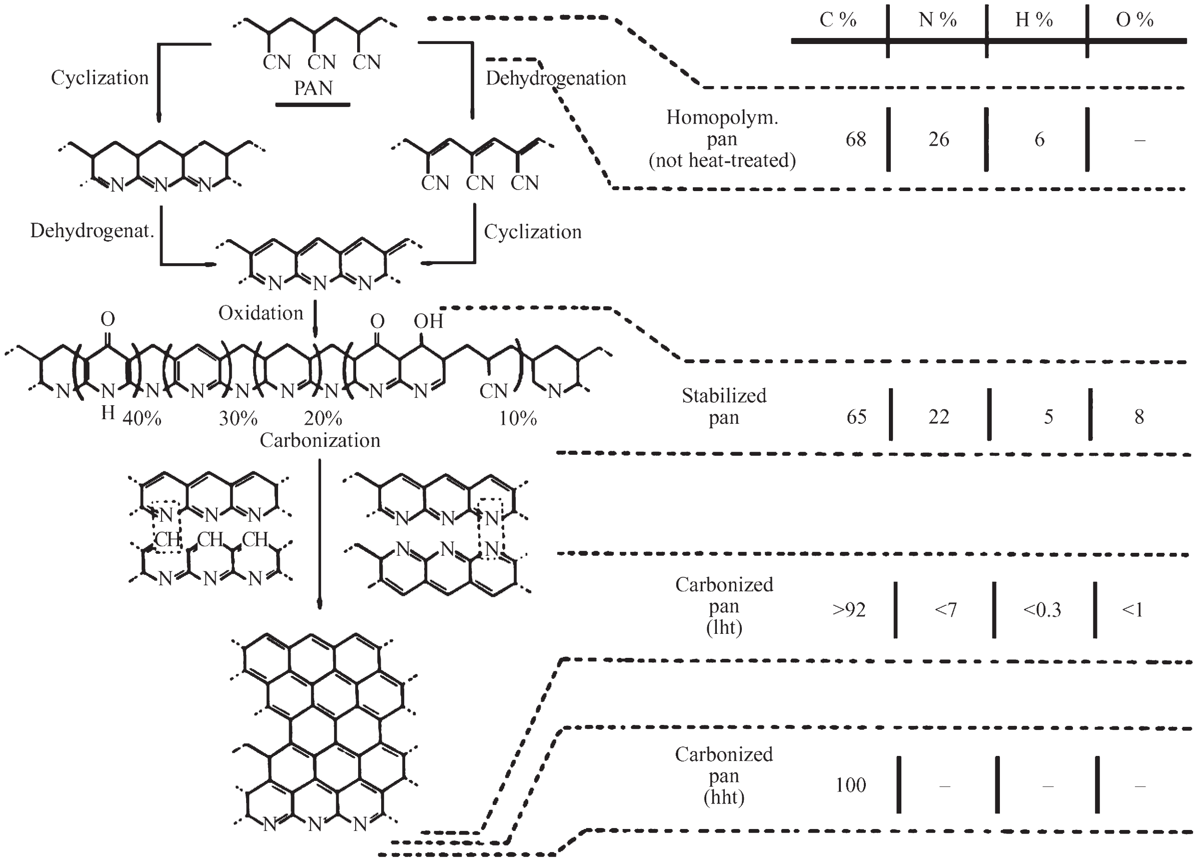

C %
N %
H %
O %
CN CN CN
PAN
Cyclization
Dehydrogenation
Homopolym. pan (not heat-treated)
68
26
6
–
CN CN CN
Dehydrogenat.
Cyclization
Oxidation
O
O OH
H 40%
30%
20%
10%
CN
Carbonization
Stabilized pan
65
22
5
8
CH CH CH
Carbonized pan (lht)
>92
<7
<0.3
<1
Carbonized pan (hht)
100
–
–
–

Figure 2
Chemical structure during processing of carbon fibers on mechanical properties of PAN.

the chemical structure of the fibers after the various process steps. From a process engineering viewpoint, mass and heat transport are the major considerations in optimizing the processing parameters.

(*b*) *Stabilization of PAN precursor fibers*. The stabilization step converts the precursor from a linear polymer to a highly condensed, thermally stable ladder polymer structure. The primary function of this step is to cross-link the as-spun structure ensuring that both molecular and the fibrillar orientations will not be lost during final heat treatment (Bahl and Manocha 1974, 1975a). The stabilization is considered to be the most decisive step in the production of carbon fibers because the stabilized fibers are considered to be the template for graphitic structures and, hence, mechanical properties of the ultimate carbon fibers.

During stabilization, the incorporation of oxygen-containing groups into the polymer chain occurs concurrently with the cyclization (Fig. 1) and dehydrogenation reactions. It is very difficult to know the exact succession of the reactions because they are dependent on the precursor constituents and processing conditions but they are probably very close to each other. The cyclization reaction is highly temperature dependent with rate determined by initiation. In copolymer fibers the comonomer acid groups dominate initiation. In the absence of an initiator, the cyclization is believed to occur by a free radical mechanism whereas in the presence of an acid initiator, an ionic mechanism predominates and occurs at a lower temperature (Fitzer 1989b). This allows better control of heat evolved in copolymers than in PAN homopolymers. The controlled cyclization reaction requires uniform heating of the fiber or heat transport to, and within, the fibers. The oxidation (intake of oxygen) is a diffusion-controlled process (mass transport) and demands uniform uptake of oxygen throughout the cross-section of the fibers. During oxidative stabilization of the fibers some oxygen-containing groups (e.g., –OH, C = O,

–COOH) are formed in the backbone of the ladder polymer. These oxygen-containing groups promote cross-linking reactions in the early stages of carbonization. The rate of intake of oxygen is also influenced by the stabilization temperature. In general, during stabilization the oxygen content of the precursor fiber increases from 8% to 11% with a change in color of the fibers from white to silky black (Peebles 1995). The evolution of CO_2 due to decarboxylation reactions and HCN due to uncyclized nitrile groups causes the carbon content of the PAN fiber to decrease while the evolution of HCN and the dehydrogenation reaction evolving H_2O cause the hydrogen content to decrease. Overall, during stabilization, the fiber suffers a net weight loss of ~5–8%. The fiber shrinks in the longitudinal direction by ~10–18% and in the radial direction by ~40–45%. This results in close packing of the structure and an increase in density of ~20–25%. Ladder polymer formation and cross-linking through oxygen are both exothermic processes. Taking these into consideration, oxidation stabilization has to be carried out at a carefully controlled heating rate and end temperature. In most commercial processes, the stabilization of PAN precursor fiber is done at a heating rate of 2–4 °C min^{-1} at temperatures ranging from 220 °C to 350 °C in the presence of air.

Alternative approaches to air stabilization have also been studied. These include the use of sulfur dioxide or bromine atmospheres, treatment of PAN fibers with cuprous chloride or potassium permanganate solution followed by air oxidation, etc. (Bahl *et al.* 1998). Though these treatments reduce the processing time, their commercial viability has yet to be established especially with respect to effluent and by-product evolution and their treatment.

Another important parameter in processing PAN-based carbon fibers is restraining the molecular chains so that they do not become disoriented during heat treatment at all stages. The chemical reactions taking place during thermal stabilization of the fibers to a ladder polymer can disrupt the orientation of molecular chains and result in linear shrinkage varying from 10–18% (Bahl and Manocha 1975b, Fitzer 1989b). In order to obtain high modulus carbon fibers it is essential to restrain the fibers from shrinking or even to elongate them by applying tension. Thus, the preferred orientation needed for the final high-performance carbon fibers is achieved by stretching the polymeric precursor fiber when it is still thermoplastic. The stretching during stabilization also results in an increase in the tensile modulus of the polymer fibers. It has been established that PAN fibers with higher elastic modulus as precursor fibers, produce carbon fibers with a higher Young's modulus. In commercial production the length change permitted may vary from −5% to +30% depending on the type of precursor fiber and its prior process history during spinning and post spinning.

(*c*) *Carbonization of stabilized PAN fibers.* The second step during the conversion of PAN fiber to carbon fiber is carbonization at temperatures of 1000–1500 °C in an inert atmosphere, usually ultrapure nitrogen. In the early stage (500–700 °C) a pyridinic structure is formed which on further heat treatment (1000–1500 °C) collapses into a turbostratic, stacked ring structure. The chemical reactions during the early carbonization stage lead to the evolution of water, ammonia, hydrogen cyanide, carbon dioxide, carbon monoxide, methane, etc. As in stabilization the sequence, nature, amount, and rate of evolution of these depend on the compositions of the precursor PAN fibers, the extent of stabilization, and the carbonization conditions. The low temperature carbonization (500–700 °C) is extremely important since it is during this stage that the maximum mass transfer takes place and the carbon building block is laid. Therefore, a low heating rate (<5 °C min^{-1}) should be adopted in this temperature range to avoid fast evolution of gaseous products and, hence, defects in the carbon fibers. In the temperature range of 700–1500 °C, H_2, HCN, and nitrogen are evolved due to intermolecular cross-linking of the polymeric chains. The amount of evolution of these gases is very low (<10%) and, hence, higher heating rates can be used in this temperature range. During carbonization the fibers suffer a weight loss of ~45–55% and shrink in diameter by ~40–50%. With heat treatment, the carbon content increases (92%), the nitrogen content decreases (7%), and the hydrogen content decreases (0.3%). The strength, as well as modulus, increase as shown in Fig. 3. Research work since 1985 on the stabilization and carbonization of PAN fibers has resulted in a many fold decrease in the processing time and an increase in mechanical properties.

Fibers heat treated to 1500 °C are termed high strength carbon fibers and consist of oriented small crystallites. Higher heat treatment (to 2500–3000 °C) continues to improve the graphitic structure with an increase in crystallite size, preferred orientation, and, hence, elastic modulus. The strength of the fiber, however, decreases, reaching a minimum at about 1800 °C and a slight improvement after heat treatment of 2000 °C (Fig. 3). This is due to the development of flaws as a result of the evolution of nitrogen and lattice rearrangement and defect removal at high temperatures. The carbon fibers heat treated in this temperature range consist of more than 99% carbon and are termed *high modulus carbon fibers.*

(*d*) *Structure of PAN-based carbon fibers.* The structure and structure-property relationships of various ex-PAN carbon fibers have been studied by many scientists (Diefendorf and Tokarsky 1975, Johnson 1987, Guigon *et al.* 1984). The structure of PAN-based carbon fibers is fibrillar in nature, similar to

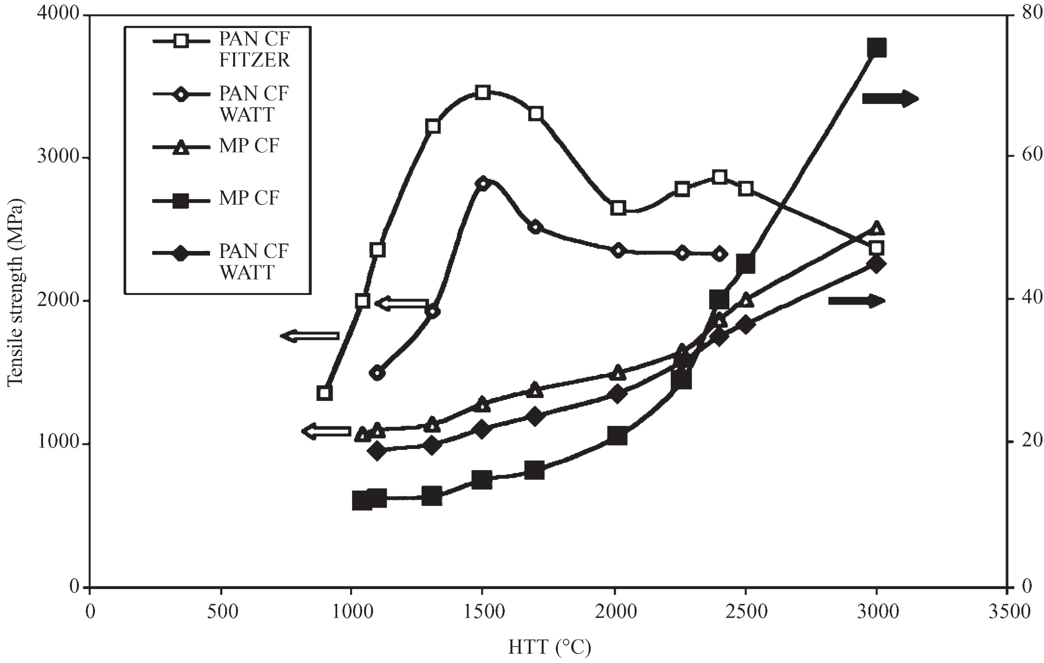


Figure 3
Influence of heat treatment on mechanical properties of PAN and pitch-based carbon fibers (after Edie 1998).

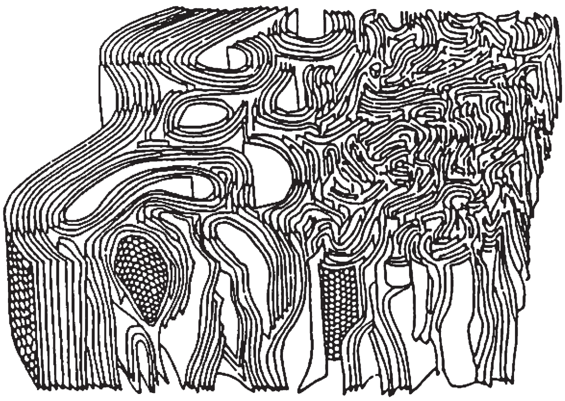


Figure 4
Johnson's (1987) model for the structure of HM PAN based carbon fibers.

that of the original precursor fiber but consisting of layers of carbon. A 3D schematic representation of the microstructure of PAN-based carbon fibers given by Johnson (1987), is given in Fig. 4. The structure of high-modulus carbon fibers consists of extensively folded and turbostratic layers of carbon with an interlayer spacing of 0.34–0.37 nm, considerably larger than that of pure graphite and crystallite size $L_a \sim 16$–23 nm and $L_c \sim 6.2$–7.0 nm, and much smaller than those of pure graphite. The structure of the fibers varies from surface to the center and is thus termed a skin-core structure, possessing no regular 3D order. In the outer region of the fibers, the turbostratic layers within the fibrils tend to be oriented parallel to fiber axis but contain needle shaped voids within the crystallites. In the core region the layer planes are folded extensively. Guigon *et al.* (1984) proposed the structure of PAN-based carbon fibers consisting of wrinkled aromatic carbon layer sheets called basic structural units (bsu) which are then crumpled along the fiber axis and are folded and entangled across the section. This disordered fibrillar structure of ex-PAN carbon fibers makes it less prone to flaw-induced failure.

1.3 Pitch-based Carbon Fibers

(*a*) *Pitch precursor and fiber spinning*. Pitch, because of its high carbon content and yield, is the prime precursor for more than 80% of industrial carbon products. Being a cheap raw material, it can bring down the cost of carbon fibers and widen the application of these materials. Therefore, concurrently with the development of PAN-based carbon fibers, efforts have also been made to develop pitch-based carbon fibers. Otani (1965) developed the first pitch-based carbon fiber from isotropic pitch. The mechanical properties were not very attractive and the fibers were termed general purpose carbon fibers. Subsequently they made carbon fibers from mesophase pitch possessing a high modulus (Otani *et al.* 1972). Like rayon-based carbon fibers, pitch-based carbon fibers could also be hot stretched in the temperature range 2000–2800 °C resulting in a two to three fold increase in Young's modulus. But commercial exploitation of the process is difficult. Barr *et al.* (1976) at Union Carbide found that an oriented fibrous microstructure could be obtained during spinning while the pitch is in the liquid crystalline state. Carbon fibers produced by this process possessed very high modulus. Therefore, Union Carbide was the first to commercialize pitch-based, high-performance carbon fibers. Since then there has been continuing interest throughout the world, especially in the US and Japan, in producing pitch-based carbon fibers with Young's modulus and thermal conductivity values approaching those of single crystal graphite (Singer 1989). While most of the PAN-based carbon fiber researchers have focused on oxidation and high temperature treatment of PAN fibers, most pitch-based fiber scientists have concentrated on pitch precursor chemistry, development of suitable mesophase pitch, and mesophase pitch spinning.

Pitches consist of a wide range of polyaromatic and cyclic aliphatic hydrocarbons containing linear aliphatic side chains. Molecular weights of the polyaromatic compounds vary from 200 to 1000. Because of this variation in molecular composition, the melting of pitches from natural sources such as coal tar and petroleum occur over a wide temperature range. When a highly aromatic pitch is heated to 400–500 °C for extended periods, called thermal polymerization, a liquid crystalline material consisting of large polynuclear aromatic hydrocarbons, a mesophase, is formed. Following its discovery by Brooks and Taylor (1965), the formation of mesophase, mode of growth, and its properties have been the subject of much research. The formation reaction and also the extent of mesophase formation depend on the carbon to hydrogen ratio of the aromatic pitches as well as the heat treatment temperature. In addition to the production of mesophase from natural sources such as coal tar and petroleum pitches, significant success has been achieved in producing synthetic mesophases with improved uniformity and purity from pure aromatic hydrocarbons (Mochida *et al.* 1995). The first commercial mesophase precursors were produced by Union Carbide using a thermal polymerization process. In addition to thermal polymerization techniques, solvent extraction, supercritical fluid extraction (Bolanos and Thies 1996), and catalytic polymerization techniques have been explored to produce mesophase pitches. The mesophase products produced by these techniques differ considerably in terms of molecular weight distribution and concentration of aliphatic side chains on the mesophase molecules. The aromatic molecules in the liquid crystalline mesophase can be depicted as a collection of slippery sheets. These can be easily extended and drawn into highly oriented filaments or fibers by shear or elongational forces applied above the melting temperature of the mesophase (Edie 1998). In order to reduce the viscosity of the pitch and the spinning temperature, which would allow for easier processing, a small amount of isotropic pitch could be added to the mesophase pitch. However, due to the solubility of the two phases in one another and interactions on heating, the melt spinning of pitches containing less than 100% mesophase is extremely difficult. There is also a stability problem encountered with two-phase extrusion.

The mesophase pitches used to produce carbon fibers are normally a mixture of high molecular weight molecules containing a small number of side groups. Both of these characteristics are vital for the melt spinning of pitch. Mesophase pitches of high molecular weight compounds with no side groups, or containing compounds with a high specific softening point, often decompose before flowing and, hence, are not preferred for high performance carbon fibers.

In contrast to PAN or rayon-based carbon fiber processing, in the pitch route, the pitch has to be first spun into fibers. The type and composition of pitch yielding ultimate carbon fibers with desired properties can be selected for making pitch fibers and even the melt spinning conditions can be controlled and regulated. Moreover, the transverse microstructure and necessary axial orientation in pitch-based carbon fibers is achieved during melt spinning of the liquid crystalline mesophase. These are maintained during stabilization and are further improved during carbonization and high-temperature heat treatment. Therefore, the pitch route offers significant advantages over the other processes. The viscosity of mesophase pitches, and hence, their flow characteristics are extremely temperature dependent. Figure 5 (Edie 1998) shows a schematic diagram of a typical melt spinning process used to produce mesophase pitch precursor fibers. The molten pitch precursor, after being transferred to a metering pump by extruders, is forced into a spin pack consisting of a filter and a spinneret. As fibers are drawn through the spinneret, they are simultaneously quenched by the surrounding atmosphere. This also induces stresses in the fibers, which in turn affect the ultimate microstructure. In order to obtain the desired structure and mechanical properties of the end carbon fibers, it is important to have precise control of the spinning conditions.

(*b*) *Stabilization of pitch fibers*. The second step in pitch carbon fiber processing is the thermostabilization of pitch fibers. Stabilization before carbonization is necessary to prevent the fibers from melting or relaxing during the final heat treatment. This can be accomplished either by heating the pitch fiber to a

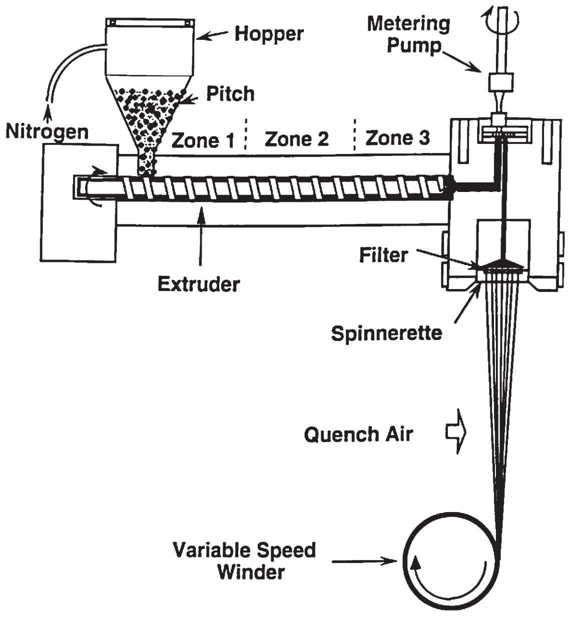


Figure 5
Schematic melt spinning process to produce mesophase pitch fibers.

temperature below its softening point in the presence of oxygen for an extended period or by immersing the precursor fiber in a strong oxidizing liquid. During this step, the large plate-like molecules become cross-linked through carbonyl and phenoxy groups formed by oxygen. According to Mochida *et al.* (1995) the stabilization reaction first occurs at the aliphatic side chains of the mesogen molecules with the introduction of oxygen functional groups. This is followed by aromatization and cross-linking. The reactivity of a mesophase pitch fiber depends on the amount of naphthenic and alkyl groups present in the material. The weight of the fibers is also increased by 5–10%. Like PAN fibers, the oxidative stabilization of pitch fibers involves simultaneous diffusion of oxygen and chemical reaction.

Both phenomena are temperature and time dependent and have to be optimized from both a processing point of view as well as the fiber properties. Typically, the pitch fibers are stabilized in the temperature range 250–300 °C for a period of 2–5 hours. Often the stabilization temperature begins near the softening point of the mesophase and is increased in a series of steps to the end temperature. Isotropic pitch fibers are stabilized at higher temperatures and take longer to stabilize than anisotropic pitch fibers. Also the oxygen uptake is greater in the former than in the latter. There is a difference of opinion about the need to stretch the pitch fibers during stabilization. However, since the as-spun fibers are already oriented, tension is not applied during stabilization in the commercial process. Because the pitch fibers have very low strength and are brittle, which poses handling problems, the stabilization is carried out on conveyer belts or on bobbins in batches.

(*c*) *Carbonization and high temperature heat treatment of pitch fibers.* The stabilized pitch fibers are further heat treated to 1500–3000 °C in an inert atmosphere during which most of the noncarbon elements escape in the form of hydrogen, carbon monoxide, methane, etc. leaving behind carbon in a well-oriented form. Carbonization or heat treatment of pitch fibers is carried out in steps. First the fibers are heat treated at ~700 °C wherein not only the carbon content in the fibers increases but they also attain flexibility (5–8%) making them easier to handle. They are further heat treated to the desired temperature of 1600–2800 °C. Since pitch-based fibers are rich in carbon content, the yield of the fibers during carbonization is very high (70–80%), and the cross-sectional shrinkage is low (25–30%). The graphitic content, as well as the degree of preferred orientation of the crystallites within the fiber, increases with heat treatment temperature.

(*d*) *Structure of pitch-based carbon fibers.* Pitch-based carbon fibers primarily consist of axially oriented graphene layer planes having different transverse microstructures. The transverse texture of most commercial pitch carbon fibers are either random, onion skin, flat layer, radial, or radial folded as shown in Fig. 6 (Fitzer and Manocha 1998). These textures are developed during mesophase spinning and are retained during subsequent stabilization and heat treatment steps. Different textures can be obtained either by disrupting the flow during extrusion, varying extruder geometry, or through variation of temperature and flow during spinning (Hamada *et al.* 1987). According to Rey (1995), a high temperature of extrusion would favor the formation of onion skin textures, low temperature would favor radial textures while intermediate temperatures result in random textures in the ultimate carbon fibers. By using stirrers of different shapes, pitch fibers with random, onion, and quasi-onion shaped microstructures can be obtained. Edie *et al.* (1994) used rectangular extrusion capillaries to develop ribbon shaped fibers with a linear transverse texture (Fig. 7). The microstructure of pitch-based carbon fibers may be considered as consisting of plates made of graphene layers oriented at angles of 2–15° to the fiber axis. The plates may be flat or constitute folds with crystallite dimensions, L_a and L_c, of 54–87 nm and 13–30 nm, respectively and interlayer spacing (d_{002}) of 0.338–0.340 nm. This plate-like structure of the mesophase pitch carbon fibers makes it more flaw

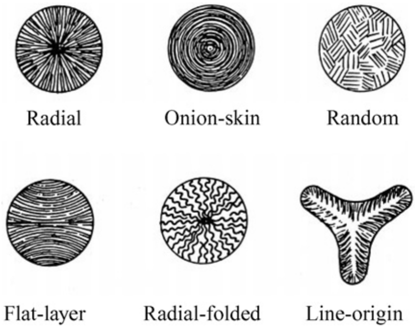


Figure 6
Transverse microstructure of pitch-based carbon fibers.

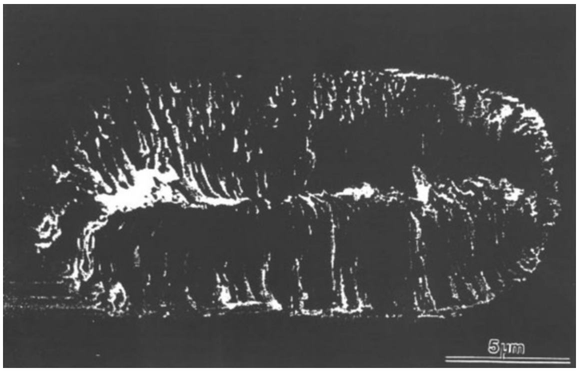


Figure 7
Transverse texture of ribbon shape carbon fiber extruder from rectangular capillary with converging entry (after Edie 1998).

sensitive. Mesophase pitch carbon fibers with random or radial folded textures exhibit higher tensile and compressive strengths whereas those with a linear transverse texture develop better lattice dependent properties (modulus and conductivities).

1.4 Vapor Grown Carbon Fibers

Another variety of reinforcement with high mechanical and thermal properties and significant commercial potential is carbon fibers produced by direct growth from the vapor phase, i.e., vapor grown carbon fibers (VGCF). These are also called catalytic chemical vapor deposition (CCVD) carbon fibers. These fibers are produced by catalytic decomposition of a hydrocarbon such as benzene, methane, or propane at 1000–1500 °C in the presence of a transition metal (Fe, Ni, or Co) or metallo-organics such as ferrocene, $(C_5H_5)_2Fe$ (Tibbetts *et al.* 1987). The catalyst plays a vital role in the growth of VGCFs. The carbon fiber yield is enhanced if a catalyst with a small but broader distribution of particle sizes is used. VGCFs are produced in short lengths (50–70 mm) and small diameter (0.5–2 μm). Benzene, or other hydrocarbons which generate benzene during decomposition, are preferred as a precursor for a higher rate of production. These fibers have also been produced from low cost sources such as linz-donawitz converter gas and coal-derived hydrocarbons. These fibers possess very high mechanical properties but have a large scatter in the values.

The structure, and thus the mechanical properties of VGCF, are independent of the precursor gas source employed, but are extremely dependent on processing parameters such as temperature of growth, type and distribution of catalyst, etc. The aspect ratio of VGCF is found to be dependent on catalyst to hydrocarbon flow ratio. Low temperature deposition (<900 °C) results in vermicular filaments, while high temperatures (1500–2500 °C) favor the growth of long, straight filaments. Since VGCF are produced in a one-step process, these possess significant commercial potential as low cost carbon fibers.

Based on electron microscopic studies on the structure of the fibers, the growth process has been proposed consisting of two parts (Endo *et al.* 1977). In the primary process a thin tube of carbon is formed by catalytic growth on a catalyst particle. This is followed by secondary growth onto the first tube. VGCF consists of turbostratic carbon layers parallel to the fiber axis arranged in annular concentric sheets like rings in a tree trunk. The core is more perfectly ordered while the secondary sheets may contain some defects. VGCF are graphitizable carbons, and when heat treated to more than 2500 °C, develop a well-oriented graphitic structure and increase in density and Young's modulus.

2. *Properties of Carbon Fibers*

The mechanical properties of ex-PAN and ex-mesophase pitch (MP) carbon fibers are dependent on the heat treatment temperature of the fibers. The variation of mechanical properties of MP carbon fibers with heat treatment temperature is also included in Fig. 3. The tensile strength of PAN-based carbon fibers goes through maxima at about 1500 °C and 2000 °C, whereas the tensile strength of mesophase pitch-based carbon fibers continuously increases with heat treatment temperature. However, it is always lower than that of ex-PAN carbon fibers. The Young's modulus of both types of fiber increases continuously with heat treatment temperature. PAN-based carbon fibers heat treated to 1500 °C are termed high strength (HT) fibers, those heat treated to ~2000 °C are termed high strength high modulus (IM) fibers, and those heat treated to 2500 °C and above are termed high modulus (HM) fibers. Mesophase pitch (MP) carbon fibers consisting of

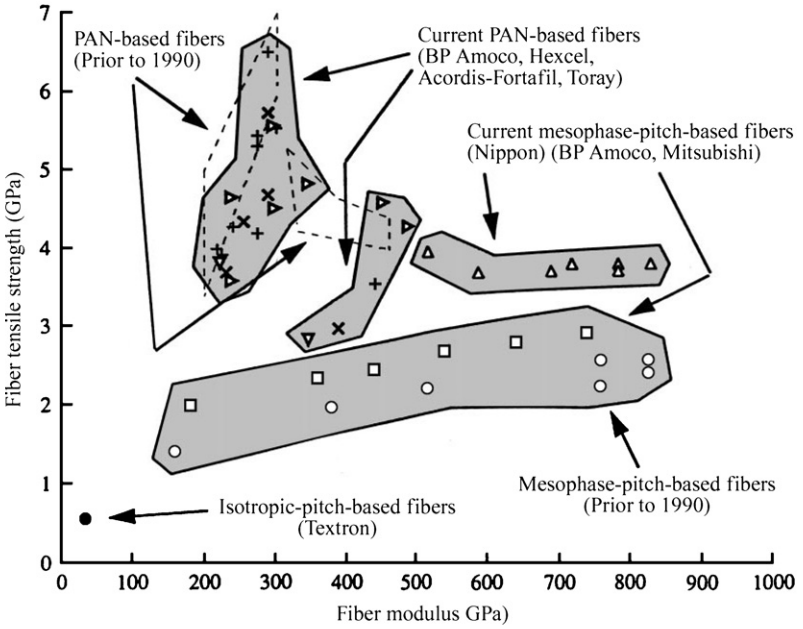


Figure 8
Present status of mechanical properties of commercial PAN-based and MP-based carbon fibers (after Edie 1998).

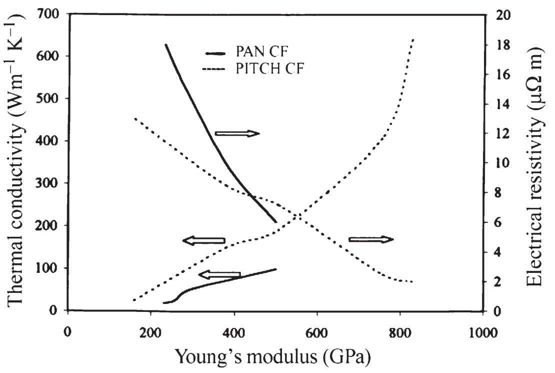


Figure 9
Young's modulus, thermal conductivity, and electrical resistance of some commercial carbon fibers.

well packed highly oriented graphitic layers posses very high densities ($>2.0\,\mathrm{g\,cm^{-3}}$) and a Young's modulus close to that of perfect graphite. MP carbon fibers possess a higher Young's modulus (600–700 GPa) than PAN-based carbon fibers. Mechanical properties of some of the commercial carbon fibers are shown in Fig. 8. Although VGCF posses a highly oriented graphitized structure, their mechanical properties are in the same range as for PAN-based carbon fibers. The thermal properties of carbon fibers are also controlled by the size, orientation, and perfection of crystalline regions. Since Young's modulus is known to be directly related to the microstructural parameters, correlation is often derived between Young's modulus and the thermal properties of various carbon fibers. Figure 9 shows the correlation between Young's modulus and thermal conductivity of various carbon fibers. In general, PAN-based carbon fibers exhibit lower thermal conductivity than pitch-based carbon fibers. High performance MP pitch-based carbon fibers with ordered microstructure exhibit very high thermal conductivity nearing $1000\,\mathrm{W\,m^{-1}K^{-1}}$. These are known as high thermal conductivity carbon fibers. VGCFs with much higher thermal conductivity, in the order of $2000\,\mathrm{W\,m^{-1}K^{-1}}$, have been developed on a pilot scale. Mesophase pitch-based HM carbon fibers also possess excellent electrical and electronic properties.

3. *Concluding Remarks*

Carbon fibers are commercially available in different forms, e.g., short fibers, continuous tows, woven and nonwoven fabrics, braids, hybrid cloths, UD tapes, and MD woven preforms. They can be combined with virtually all types of matrices, i.e., polymer, ceramic, carbon, cement, and metal to manufacture composite products for different end applications. Most applications are found with polymeric composites. Carbon fibers are given specific surface treatments and finishes compatible with a particular thermoset or thermoplastic resin system. Though initially developed for aerospace and aeronautics applications, in the 1990s these composites have been increasingly used in general industries such as sporting goods, automobiles, machinery, adsorbents and catalysts, electrical and electronic applications, and building materials. PAN-based carbon fibers, due to their balanced strength and modulus, dominate such applications whereas MP carbon fibers, with their high thermal conductivity, dominate the thermal management areas.

Isotropic pitch and rayon-based carbon fibers are excellent materials for the production of activated carbon fibers with very high surface areas ($>1500\ \mathrm{m^2\,g^{-1}}$) and are likely to dominate as materials for liquid as well as gaseous adsorption and environmental protection.

In the energy sector, VGCF and carbon fibers with a platelet structure exhibit excellent hydrogen storage capacity. Carbon fibers can act both as a conductor, like metals, as well as a semiconductor. Therefore, these materials open up the possibility of their use in smart structures. Carbon fibers are known to have excellent biocompatibility and as such, as well as in composite form, have been successfully used as bioimplants (see *Carbon in Biomedical Engineering*[#]). Carbon fibers have low noise characteristics and have been successfully used in neurobiology for electrochemical detection of released compounds from single brain cells. Newer applications are being explored in various fields.

Future efforts are likely to concentrate on producing PAN-based carbon fibers with much higher strength and moduli, approaching those of pure graphite. General purpose, low cost, noncircular fibers with balanced mechanical properties, in addition to high thermal conductivity, will continue to be an attraction in the pitch sector.

See also: Carbon Nanotubes*

Bibliography

Bacon R 1973 Carbon fibers from rayon precursor. In: Walker P L, Thrower P A (eds.) 1973 *Chemistry and Physica of Carbon*. Dekker, New York, pp. 1–102

Bacon R, Schalamon W A 1969 Physical properties of high modulus graphite fibers made from a rayon precursor. *Appl. Polym. Symp.* **9**, 285–92

Bahl O P, Manocha L M 1974 Characterization of oxidised PAN fibers. *Carbon* **12**, 417–23

Bahl O P, Manocha L M 1975a Effect of preoxidation conditions on the mechanical properties of carbon fibers. *Carbon* **13**, 297–300

Bahl O P, Manocha L M 1975b Shrinkage behaviour of polyacrylonitrile fibers during thermal treatment. *Angew. Makromol. Chem.* **48**, 145–59

Bahl O P, Shen Z, Lavin J G, Ross R A 1998 Manufacture of carbon fibers. In: Donnet J B, Want T K, Peng J C M (eds.) 1998 *Carbon Fibers*. Dekker, New York, pp. 1–77

Bolanos C, Thies M C 1996 Supercritical toulene–petroleum pitch mixtures: Liquid–liquid equilibria and SAFT modeling. *Fluid Phase Equilibria* **117**, 273

Barr J B, Chwastiak S, Didchenko R, Lewis I C, Lewis R T, Singer L S 1976 High modulus carbon fibers from pitch precursor. *Appl. Polym. Symp.* **29**, 161–73

Brooks J D, Taylor G H 1965 The formation of graphitizing carbons form the liquid phase. *Carbon* **3**, 185–93

Diefendorf R J, Tokarsky E 1975 High performance carbon fibers. *Polym. Eng. Sci.* **25**, 150–9

Donnet J P, Bansal R C 1984 *Carbon Fibers*. Dekker, New York

Edie D D 1998 The effect of processing on the structure and properties of carbon fibers. *Carbon* **36**, 345–62

Edie D D, Diefendorf R J 1993 Carbon fiber manufacturing. Buckley J D, Edie D D (eds.) *Carbon/carbon Materials and Composites*. Noyes, Park Ridge NJ, pp. 19–37

Edie D D, Robinson K E, Fleurot O, Jones S P, Flain C C 1994 High thermal conductivity ribbon fibers from naphthalene-based mesophase. *Carbon* **32**, 1045–54

Endo K, Oberlin A, Koyama T 1977 High resolution electron microscopy of graphitizable carbon fiber prepared by benzene decomposition. *Jpn. J. Appl. Phys.* **16**, 1519–23

Fitzer E 1985 *Carbon Fibers and their Composites*. Springer, Berlin

Fitzer E 1989a PAN based carbon fibers—present state and trend of the technology from the viewpoint of possibilities and limits to influence and to control the fiber properties by the process parameters. *Carbon* **27**, 621–45

Fitzer E 1989 Carbon fibers—present status and future expectations. In: Figueiredo J L, Bernardo C A, Bakar R T K, Huttinger K J (eds.) *Carbon Fibers Filaments and Composites*. Springer, New York, pp. 3–41

Fitzer E, Heine M 1988 Carbon fiber manufacturing and surface treatment. In: Bunsell A (ed.) *Fiber Reinforcements for Composite Materials*. Elsevier, Amsterdam

Fitzer E, Manocha L M 1998 *Carbon Reinforcements and Carbon–Carbon Composites*. Springer, Berlin

Guigon M, Oberlin A, Desarmot G 1984 Microtexture and structure of some high-modulus PAN-based carbon fibers. *Fiber Sci. Technol.* **20**, 177–98

Hamada T, Nishida T, Sajiki Y, Matsumoto M, Endo M 1987 Structures and physical properties of carbon fibers from coal tar mesophase pitch. *J. Mater. Res.* **2**, 850–7

Johnson D J 1987 Structure property relationships in carbon fibers. *J. Phys. D: Appl. Phys.* **20**, 287–91

Johnson W, Watt W, Phillips L N, Moreton R 1966 *Br. Pat.* **1** 166–251

Mochida I, Yoon S-H, Korai Y, Kanno K, Sakai Y, Komatsu M 1995 Carbon fibers from aromatic hydrocarbons. *Chemtech.* **25**, 29–37

Otani S 1965 On the carbon fiber from the molten pyrolysis product. *Carbon* **3**, 31–8

Otani S, Watanabe S, Ogino H, Iijima K, Koitabashi T 1972 High modulus carbon fibers from pitch materials. *Bull. Chem. Soc. Jpn.* **45**, 3710–4

Peebles L H 1995 *Carbon Fibers: Formation, Structure and Properties*. CRC, Boca Raton, FL

Rajalingam P, Radhakrishnan G 1991 Polyacrylonitrile precursor for carbon fibers. *J. Macromol. Soc. Rev. Macromol. Chem. Phys. C* **31**, 301

Rey A D 1995 Elasticity-driven texture selection mechanism in mesophase carbon fibers. *Phys. Rev. E.* **52**, 6278–81

Shindo A 1961 Studies on graphite fibers. Report No. 317. Government Industrial Research Institute, Osaka, Japan

Singer L S 1989 Carbon fibers. In: Kelly A (ed.) *Concise Encyclopedia of Composite Materials* Pergamon Press, Oxford, UK

Tibbetts G G, Devour M G, Rodda E G 1987 An adsorption diffusion isotherm and its application to the growth of carbon filaments on iron catalyst particles. *Carbon* **25**, 367–75

Watt W, Johnson W 1969 The effect of length changes during oxidation of polyacrylonitrile fibers on the Young's modulus of carbon fibers. *Appl. Polym. Symp.* **9**, 215–27

L. M. Manocha
Sardar Patel University, Gujarat, India

Carbon Nanofibers

Considerable interest has been shown in understanding the process mechanism for the formation of carbon nanofibers, also commonly referred to as filamentous carbon. Carbon nanofibers are produced from the catalytic decomposition of hydrocarbon gases or carbon monoxide over selected metal particles that include iron, cobalt, nickel, and some of their alloys at temperatures over the range 400–1000 °C. Initial investigations in this area were motivated by the need to inhibit the growth of carbon nanofibers because of the persistent problems caused by accumulation of the material in a variety of commercial processes. Fundamental studies have demonstrated that carbon nanofibers possess an unusual combination of properties that include high surface area and high electrical conductivity. It is these features that have stimulated research efforts that are designed to optimize and control the formation of this type of carbon. One of the most outstanding characteristics of highly graphitic nanofibers is the presence of a large number of edges, which provide a tailored surface consisting of sites that are readily available for interaction with gases, liquids, or other solids. In this article the key steps in the growth mechanism of carbon nanofibers will be discussed, with emphasis being placed on the factors that control the degree of crystalline perfection and molecular architecture that can be achieved with these materials. Attention is also focused on the unique set of physical and chemical properties exhibited by carbon nanofibers, and areas where the material is finding application.

1. Growth Mechanism of Carbon Nanofibers

The growth mechanism leading to the formation of carbon nanofibers has been studied by many different groups and has been a source of debate, particularly with regard to the chemical state of the catalyst species and the driving force for the reaction. *In situ* electron microscopy techniques were used to observe directly the manner by which small metal particles generated carbon nanofibers during the decomposition of acetylene (Baker *et al.* 1972). From analysis of the recorded videotape sequences it was possible to measure rates of growth of the material and determine some of the kinetic parameters involved in the process. On the basis of these experiments, a growth mechanism was proposed that was later refined to include the following steps: (i) adsorption and decomposition of the reactant hydrocarbon molecule at a certain set of metal crystallographic faces, (ii) dissolution and diffusion of carbon species through the metal particle, and (iii) precipitation of carbon at a different set of crystallographic faces to form the nanofiber structure. It was argued that diffusion of carbon through the metal catalyst particle was the rate-determining step. This claim was substantiated by the close agreement between the measured activation energies for nanofiber growth and that for carbon diffusion through the respective metals that were used as catalysts in the process.

The chemical nature of the metal catalyst, the reaction temperature, and the composition of the reactant gas dictates the morphology and degree of crystalline perfection exhibited by the carbon nanofibers. Figure 1 is a schematic diagram designed to illustrate the key features of the growth mechanism

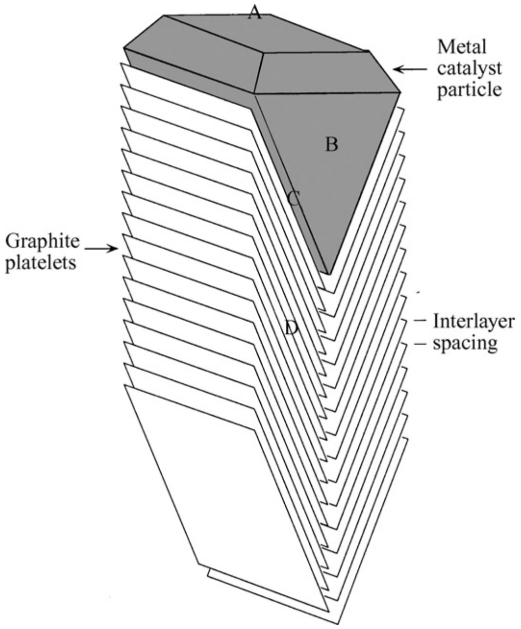


Figure 1
Schematic rendition of a graphite nanofiber and associated catalyst particle highlighting the key steps in the growth process.

for a particular type of carbon nanofiber structure. The hydrocarbon gas or carbon monoxide is initially absorbed and decomposed on the metal face (A) of the catalyst particle. This reaction with the surface results in the formation of chemisorbed carbon species that proceed to dissolve and diffuse through the bulk of the metal particle (B), ultimately leading to precipitation of carbon in the form of graphite platelets at the rear of the metal particle (C). The nanostructure formed in this example adopts a "herringbone" stacking arrangement (D) of the graphite platelets. As the heat balance is maintained in the reaction system, the carbon structure will continue to grow until perturbations occur in the chemistry of the leading face of the metal catalyst, thus causing a buildup of solid carbon at this interface, which results in deactivation and culmination in nanofiber growth. The driving force responsible for the carbon diffusion step is still not well understood. In the initial studies it was hypothesized that a temperature gradient was established within the catalyst particle, which resulted from an exothermic hydrocarbon decomposition reaction at the metal/gas interface and an endothermic precipitation reaction at the metal/solid carbon interface. An alternative and somewhat more plausible rationale has been suggested that relies on the participation of a carbon concentration gradient across the particle (Alstrup 1998).

One of the ramifications of the nanofiber growth process is that these structures can provide a means of transport for small metal particles. Initially the catalytic entity may be located on a support medium, at the grain boundary of a bulk piece of metal or part of a large powdered granule. As the growth of the nanofiber proceeds, the catalyst particle is detached from its original location and carried away from such a surface since it is now attached to part of the solid carbon structure. This aspect can have serious repercussions with regard to the behavior of a heterogeneous catalyst system, since the nature of the interaction with the support is being radically changed from the original state to one where the metal is exclusively in contact with carbon. This type of perturbation can cause major modifications in the performance of the catalyst system and in certain circumstances can lead to premature deactivation.

2. *Synthesis of Carbon Nanofibers*

Carbon nanofibers can be prepared in relatively large quantities using a variety of metals either in a powdered or supported form as the catalytic entities. The following protocol is followed for the synthesis of this form of carbon by the research group at Northeastern University. The metal catalyst is dispersed in a uniform manner across the bottom in a ceramic boat that is positioned in the center of a horizontal tube furnace. Initially, a 10% hydrogen/helium gas mixture is introduced into the system to reduce the sample at 600 °C and after an appropriate interval of time the temperature is quickly brought to the desired reaction condition. A predetermined hydrocarbon/hydrogen or carbon monoxide/hydrogen reactant mixture is introduced into the system and the reaction generally allowed to proceed for periods of up to two hours. It is essential to regulate the supply of the various reactant gases by using well-calibrated mass flow controllers. Immediately following the growth sequence the reactant gas is replaced with helium and the sample cooled to room temperature. This is the general method for growing the carbon nanofibers in the laboratory and can produce various quantities of fibers; the most active systems can produce as much as 15 g quantities from a 100 mg catalyst charge weight.

In addition to powdered metals, other catalyst configurations are equally effective for the growth of carbon nanofibers: these include gauzes, foils, wires, or supported metal particles. These approaches can offer a simpler method for handling the nanofibers for various applications. When the nanofibers are generated from a supported metal catalyst system, it is of paramount importance to take into consideration the effect of the metal–support interaction on the subsequent mode of growth of the solid carbon. The nature of the support material will also have a direct impact on the location and size of the metal catalyst particles. These factors will in turn dictate the morphology and the properties exhibited by the nanofibers produced when the catalyst system is reacted in a hydrocarbon environment. When preparing a supported catalyst, the goal is to achieve a uniform distribution of the metal particles on the support medium. The metal particles for a given metal can be introduced onto the support surface using various methods and these are based on either aqueous or nonaqueous impregnation with solutions of metal salts or by evaporation of organic compounds that contain the metal particles. Indeed, if the aim is to optimize and engineer the growth characteristics of the nanofibers, this can be most easily accomplished by the careful choice of the support medium.

3. *The Events Occurring at the Metal–Gas Interface*

One of the major factors associated with the initiation of nanofiber growth is the nature of the interaction between the metal catalyst particle and the carbon-containing reactant gas. In order for the reaction to proceed, the hydrocarbon must undergo dissociative chemisorption on the metal surface, a process that leads to rupture of C=C and C—H bonds in the molecule. In many cases, this step requires a rearrangement of the metal atoms in the set of crystallographic faces at which adsorption of gas-phase species occurs. Since hydrogen is known to facilitate the reconstruction of metal surfaces, the presence of this gas in the reactant mixture plays a key role in the carbon nanofiber growth and also has a direct impact on the catalyst lifetime. It is important to note that even though a given series of metals have similar surface orientations, the characteristics exhibited by the faces do not exhibit the same pattern of behavior with respect to adsorption and decomposition of hydrocarbons. Various options are available to increase the catalytic activity of a given metal towards the growth of carbon nanofibers that are based on the introduction of other elements either directly into the catalyst particles or as additives to the gaseous reactant mixtures. These two aspects will be discussed under separate headings.

3.1 Incorporation of Solid Additives

The introduction of other metallic components into a host metal has been found to exert a dramatic impact on the behavior of the catalytic system towards the growth of carbon nanofibers. It is frequently found that when copper is added to one of the iron-type transition metals there is an enhancement in the formation of the solid carbon product and furthermore, the morphology and degree of crystalline perfection of the nanofibers is altered to a significant extent. Perhaps one of the most intriguing reactions was observed with the copper–iron system. When copper or iron powders were heated in the presence of an ethylene/hydrogen mixture at 600 °C the metals exhibited little or no ability to catalyze the formation of carbon nanofibers. In sharp contrast, when a very small amount of one of these metals was added to the other component then, on subsequent exposure to the same reactant gas, prolific growth of this form of solid carbon was achieved (Krishnankutty *et al.* 1996). In other investigations it was found that if as little as 1.0 wt.% of either silver or tin (Chambers *et al.* 1996, Nemes *et al.* 1998) was added to powdered cobalt, these additives also caused a dramatic increase in the activity of the host metal towards the production of carbon nanofibers from the decomposition of ethylene. It was suggested that some of the observed enhancement in catalytic activity could be due to the existence of electronic perturbations of cobalt, induced by the presence of either tin or silver, that are manifested in modifications in the adsorption characteristics and subsequent decomposition characteristics of ethylene on the bimetallic surface.

3.2 Incorporation of Gas-phase Additives

In investigating the various features of carbon nanofiber growth one thrust of the research has focused on the effects of incorporating controlled amounts of certain nonmetallic species into the catalyst system via the addition of such species to the gaseous reactant mixture. Studies indicate that the activity of the metal catalyst towards carbon nanofiber formation can be promoted to a significant degree by introducing controlled amounts of certain gas-phase additives such as hydrogen, sulfur, carbon monoxide, and chlorine (Kim *et al.* 1993, Rodriguez *et al.* 1993, Chambers and Baker 1997).

Treatment of iron in an ethylene/hydrogen mixture at 600 °C was found to result in only minor decomposition of the hydrocarbon and, as a consequence, under these conditions the formation of carbon nanofibers was insignificant. The addition of carbon monoxide to the reactant, however, caused a dramatic change in the catalytic activity of the metal particles, and under these circumstances relatively large amounts of solid carbon were produced. Maximum reactivity of iron was realized from the interaction with a C_2H_4–CO–H_2 (3:1:1) mixture. Removal of carbon monoxide from the system resulted in a rapid decline in activity; however, upon re-introduction of the additive activity was restored to its previous high level, showing that the activation–deactivation processes associated with the iron catalyst were completely reversible in nature. The pattern of behavior exhibited by the catalyst was rationalized in terms of reconstruction of the iron surface in the presence of coadsorbed carbon monoxide, a process that is responsible for the creation of faces that favor dissociative chemisorption of ethylene (Rodriguez *et al.* 1993).

In an analogous set of experiments the influence of small concentrations of chlorine (75–100 ppm) on the decomposition of ethylene over iron and cobalt catalyst systems was investigated. The results indicated that at temperatures in the vicinity of 400 °C, the catalytic activity of both metals with respect to the formation of carbon nanofibers from the decomposition of ethylene was significantly enhanced by the presence of a critical amount of chlorine in the reactant mixture (Chambers and Baker 1997). It is well known that the chemisorption of sulfur on a metal surface can result in poisoning and premature catalyst deactivation (Rostrup-Nielsen 1984). However, there are reports in the literature that very low concentrations of sulfur can actually exert a promotional effect on the performance of a metal catalyst (Reyniers and Froment 1995). In this context it was found that when an ethylene/hydrogen reactant mixture containing about 10–50 ppm levels of hydrogen sulfide was passed over iron, cobalt, or nickel at 600 °C, there was a substantial increase in the yield of carbon nanofibers compared to that found in the absence of the additive molecules.

In summary, one can conclude that the incorporation of controlled amounts of certain gas-phase (sulfur, chlorine) or solid additives (copper, tin, silver) into the reaction system can dramatically enhance the catalytic activity of the host metal towards carbon nanofiber formation. While geometric factors may play a dominant role in the modification in catalytic performance, one must also be cognizant of the fact that the introduction of the adatom can induce electronic perturbations in the surface of the host metal, which in turn will exert an impact on the adsorption characteristics of a hydrocarbon molecule.

4. Events Occurring at the Metal–Carbon Interface

The degree of crystalline perfection of the deposition carbon nanofiber is dictated by a number of factors including the chemical nature of the catalyst particle, the composition of the reactant gas, and the reaction temperature. It has been established that the graphitic nature of carbon nanofibers is controlled to a large degree by the wetting properties that exist between the metal and deposit, i.e., a strong interaction must exist between the two components. It has been demonstrated that in order for dissolved carbon to be precipitated in a highly crystalline form, there must be a registry between the interstitial spacing of the metal atoms at this interface and atomic spacing of carbon atoms in graphite (Yang and Chen 1985). Surface science studies have established that certain crystallographic faces of a metal favor precipitation of dissolved carbon in the form of graphite, whereas less ordered forms will be deposited from other faces (Goodman *et al.* 1980, Nakamura *et al.* 1989).

Investigations using controlled-atmosphere electron microscopy techniques have revealed that certain metals can readily undergo a spreading action along the edge regions of graphite under reducing conditions (Baker 1995). On the basis of these data it is possible to make predictions with regard to the structural perfection of nanofibers generated from a given catalyst system with ethylene/hydrogen at a temperature between 500 °C and 650 °C. Indeed, it is a detailed understanding of the properties exhibited by the two components at the interfacial region that enables one to catalytically engineer the growth characteristics of these novel materials.

In addition to controlling the degree of crystalline order exhibited by the nanofibers, the orientation of the crystallographic faces existing at the metal–carbon interface is responsible for imposing the particular geometric arrangement of the graphite platelets adopted by these structures (Rodriguez *et al.* 1995). Transmission electron microscopy studies have demonstrated that the nanofibers can exist in various conformations, some of which are presented in Fig. 2. In Figs. 2(a) and 2(b) the nanofibers have been generated via a bidirectional mode where the carbon precipitation process occurs from opposite faces of the catalyst particle, which remains embedded within the structure during the entire growth period. The

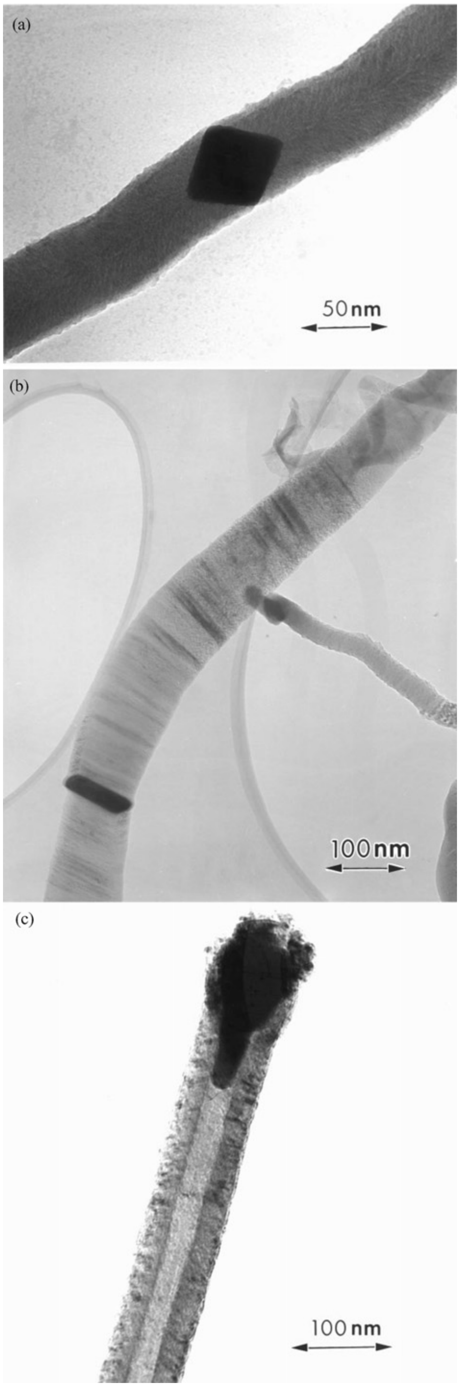


Figure 2
Transmission electron micrographs of three different types of carbon nanofibers generated by (a, b) a bidirectional mode and (c) a whisker-like process.

type of nanofiber shown in Fig. 2(c), however, is produced by a mode in which carbon precipitation occurs at the sides of the particle, which is located at the growing tip of the structure. These three types of structures share a number of common characteristics, including the fact that the width of the nanofibers is the same as that of the associated catalyst particle. A further feature is that the nanofibers are relatively straight, indicating that the carbon diffusion characteristics through the particle are symmetrical.

Figures 3 and 4 show the corresponding high-resolution micrographs and three-dimensional schematic renditions, respectively, of the carbon nanofiber structures presented in Fig. 2. Detailed examination reveals that the graphite platelets constituting the material shown in Fig. 4(b) are stacked in a direction that is perpendicular to the fiber axis, adopting the arrangement of a "deck of cards." This arrangement is to be contrasted with that presented in Fig. 4(c) where the graphite sheets are aligned parallel to the fiber axis, adopting the structure of a "multiwalled faceted nanotube." Inspection of the structure shown in Fig. 4(a) indicates that in this case the graphite platelets are at an angle to the fiber axis in a so-called "herring-bone" arrangement.

The unusual spiral form of nanofiber, highlighted in Fig. 5, is believed to originate from a perturbation in the diffusion characteristics of carbon species through the catalyst particle, brought about by the participation of either sulfur or phosphorus compounds which introduce an element of asymmetry into the system (Motojima *et al.* 1991, Owens *et al.* 1994). Such a condition could arise if the reconstruction of certain faces forced the particle to adopt an irregular shape during the formation of the nanofiber, or if adsorption of the adatoms occurred at certain faces of the particle and interfered with the carbon precipitation step, thus creating an anisotropic growth pattern.

5. *Properties of Carbon Nanofibers*

The unique configurations adopted by carbon nanofibers provide the material with a rare set of properties; a high electrical conductivity combined with a large fraction of active sites, that correspond to the exposed edge regions of the graphite platelets, are available for chemical and physical interaction with various species. Perhaps one of the most unexpected findings is that such ordered crystalline solids can exhibit high surface areas (300–700 m^2g^{-1}), where the totality of the surface area is chemically active. From the physical point of view, carbon nanofibers vary from 5 μm to 100 μm in length and are 5–100 nm in diameter.

A further factor to be taken into consideration is the functionality of the carbon nanofiber surface. When graphitic materials are in their pristine state and free of surface oxygen groups they are hydrophobic in nature. Following normal activation procedures such materials exhibit a certain degree of hydrophilic character. It has been demonstrated that following treatment in an oxygen plasma, acidic groups were introduced onto the carbon surface, while nitrogen and ammonia plasmas produce basic functionalities (Esumi *et al.* 1989). The possibility of controlling the acid–base properties of the carbon nanofiber surface could have a direct impact on a variety of potential applications.

The information gathered to evaluate the mechanical properties of carbon nanofibers is very restricted. This is because of the difficulty involved in manipulating these microscopic structures. Despite this limitation it has been established indirectly that certain types of nanofibers possess very good mechanical properties. For example, the material can be used in liquid-phase reactions where it can withstand vigorous agitation without disintegrating and also provide improved transport properties over those found with more traditional adsorbates. Furthermore, because of their size dimensions, separation of the nanofibers from liquid-phase reactants and products make such structures desirable for catalytic and filtration systems. Motojima *et al.* (1991) devised a series of experiments to determine the elastic properties of coiled carbon nanofibers by performing transmission electron microscopy to record the observations and measurements. They fixed one end of the nanostructure to a piece of copper while the other end was pulled to extend the coiled structure, thus determining the extension ratio. This parameter was determined as a function of the original coil length and it was found that the structures studied could be 3.0 times extended elastically from its original length and 4.5 times semielastically.

A technique which involved the dispersion of the material in a solvent followed by precipitation onto lithographically defined gold contacts to make a "nanowire circuit" was used to determine the electrical resistivity of individual 10 nm sized nanofibers (Whitesides and Weisbecker 1994). A resistivity of 9.5 Ωm was estimated for carbon condition along the fiber axis, which suggests that the materials could be useful for building nanoscale electronic devices.

6. *Applications of Carbon Nanofibers*

It is recognized that some exciting opportunities exist for exploiting the potential of carbon nanofibers in a variety of applications. Potential applications may include new technological devices that can possibly be used for energy storage or employed to clean the environment. For these particular applications, the surface area of the material and its adsorption properties are of great importance.

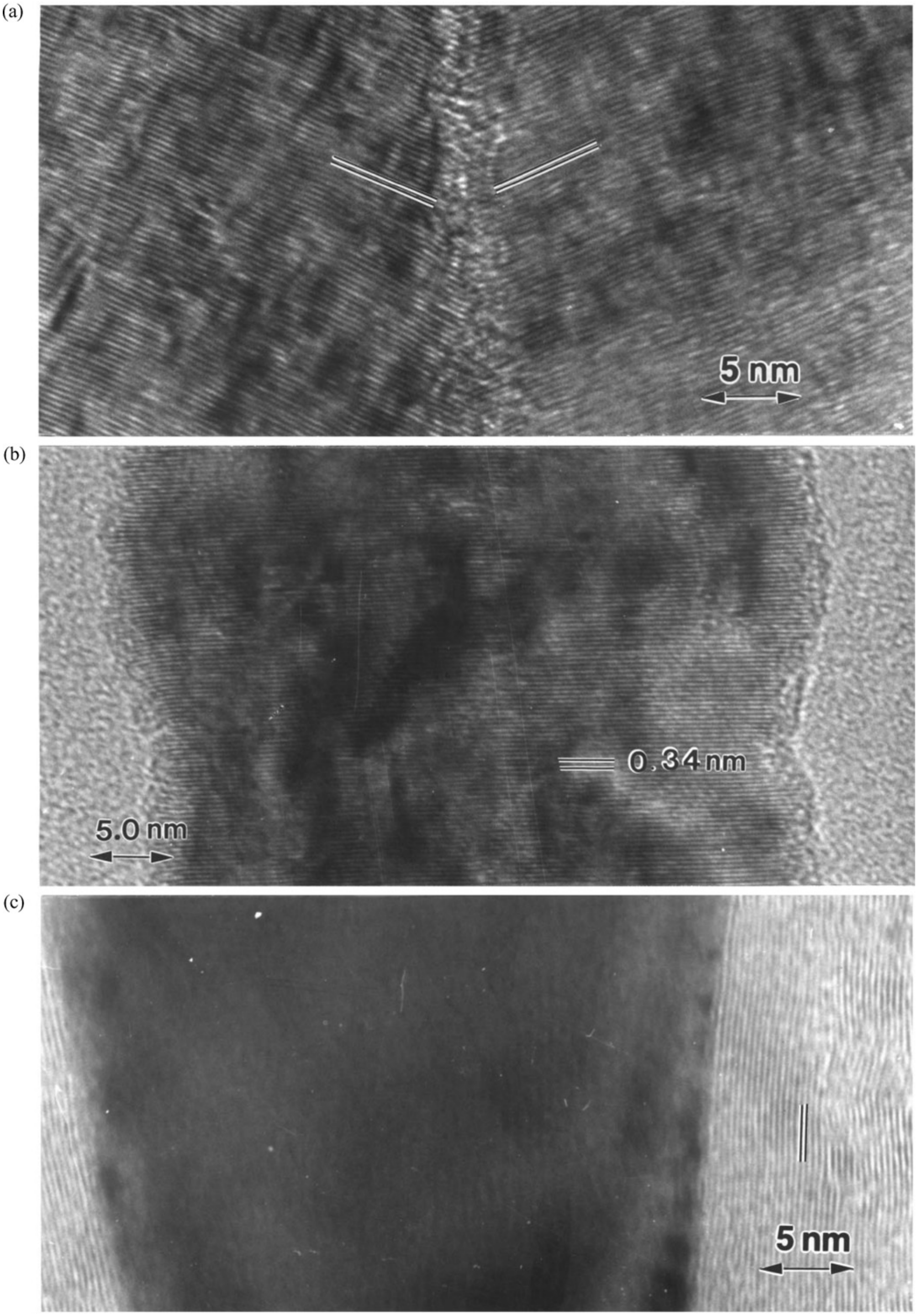

(a)
5 nm
(b)
0.34 nm
5.0 nm
(c)
5 nm

Figure 3
High-resolution electron micrographs showing the arrangements of the graphite platelets of the respective materials shown in Fig. 2.

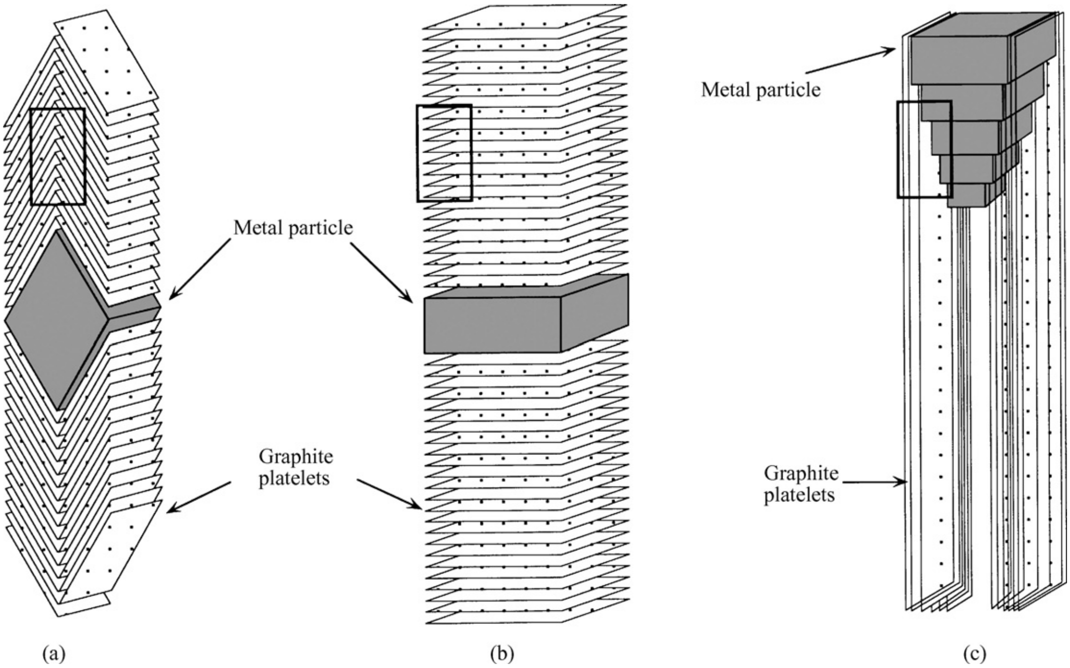


Figure 4
Schematic representation of metal catalyst particles and (a) the "herringbone", (b) the "deck of cards", and (c) the "parallel" arrangement of platelets in graphite nanofibers. Marked sections correspond to the respective electron micrographs presented in Fig. 3.

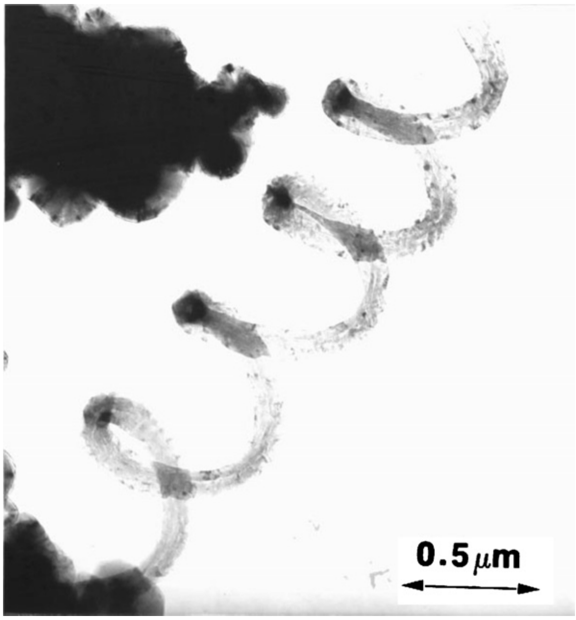


Figure 5
Transmission electron micrograph showing the spiral appearance of carbon nanofibers resulting from the interaction of Fe–Ni with ethane containing H_2S at 815 °C.

The unique blend of structural properties exhibited by certain types of carbon nanofibers, which are composed of an infinite number of extremely short and narrow pores, makes such materials ideal for trapping small organic molecules. This concept is based on the principle that a strong interaction is likely to be established between the organic impurity and the nonpolar basal regions of the graphite platelets constituting the nanofiber structure. However, adsorption of water will not be favored at these sites and, as a consequence, the material will function as a selective adsorption agent. Investigations (Fenelonov *et al.* 1995, Park *et al.* 1999) have demonstrated that this is a viable approach, from experiments in which the nanofiber adsorbate system was able to selectively remove alcohols from water in a much more efficient manner than that displayed by samples of high surface area active carbon. Furthermore, they found that the performance of the nanofibers could be improved when the materials were pretreated in 1 M hydrochloric acid.

The ability to tailor the structural characteristics of carbon nanofibers so that the product has the unusual dual properties of high surface area, normally associated with active carbon, and a high degree of crystalline perfection more typical of graphite, thus giving rise to high electrical conductivity, makes these materials ideal for use as support media. The unusual arrangement of the surface carbon atoms encountered in these materials offer some fascinating methods to control the structure of supported metal particles and induce major changes in their catalytic performance. By following this same protocol it is possible to use the nanofibers as templates for the vapor deposition of metal oxides that are likely to grow in an epitaxial format on the substrate surface. A number of investigations have focused on the modifications in both particle morphology and catalytic performance brought about by supporting metal crystallites on carbon nanofibers (Rodriguez *et al.* 1994, Planeix *et al.* 1994, Hoogenraad *et al.* 1995, Chambers *et al.* 1998a, Park and Baker 1998). The exposed surfaces of some of these materials consist entirely of graphite edges that are separated by a distance of 0.335 nm and these features provide a template for the deposition of metal crystallites where the surface atoms adopt arrangements that are not generally encountered on other conventional support media, such as active carbon, silica, or γ-alumina.

Experiments performed with nickel particles supported on graphite nanofibers showed that these systems exhibited unusual properties with regard to the selectivity patterns obtained for the hydrogenation of olefins and diolefins when compared to the behavior found when the metal was dispersed on oxide carriers (Rodriguez *et al.* 1994, Chambers *et al.* 1998a, Park and Baker 1998). This enhancement in both activity and selectivity was attributed to the fact that the nickel crystallites were located on the edge sites of the nanofibers and, as a consequence, the arrangement of the metal atoms was governed by the interaction with the carbon atoms in these regions. Under such circumstances one might expect different crystallographic faces of nickel to be exposed to the reactant gas compared to those present when the metal was dispersed on less ordered materials. This claim was supported by the observations from high-resolution transmission electron microscopy examinations which revealed the existence of major differences in the morphological characteristics of metal particles supported on graphite nanofibers and γ-alumina. Close inspection of the metal particles supported on graphite nanofibers showed that they were preferentially aligned with respect to the fiber lattice. Furthermore, the hexagonal-shaped crystallites were very thin and flat, features consistent with the existence of a strong metal–support interaction.

The use of carbon nanofiber-supported nickel catalysts was extended to a reaction that is of technological importance, namely the selective hydrogenation of α, β-unsaturated aldehydes to the corresponding unsaturated alcohol (Salman *et al.* 1999). The results of these experiments, which were carried out over the range 75–150 °C, were compared with those obtained when the same metal loading was dispersed on γ-alumina. It was extremely encouraging to find that at this upper temperature the activity and selectivity of the nanofiber-supported catalyst systems (78% crotyl alcohol) was significantly higher than the alumina-supported system (12% crotyl alcohol). This outstanding performance was attributed to a combination of factors, the unusual crystallographic orientation of nickel particles and the possibility that electronic perturbations were being induced in the metal from interactions with the electrically conductive nanofiber support media.

It is well known that advanced polymeric composites have several advantages (including high specific strength and energy absorption, low weight, styling flexibility, good noise/vibration/harshness characteristics, and excellent corrosion resistance) that make such materials ideal for fabrication of body parts for heavy vehicles. The high strength-to-weight ratio, combined with superior stiffness, has made carbon fibers the material of choice for high-performance composite structures. Carbon fibers have been used to reinforce several matrices, including polymers, metals, and ceramics. Unfortunately, these composite structures suffer from a number of shortcomings, including inadequate interlaminar shear properties and fiber pull-out. Several approaches have been used in an attempt to enhance the interfacial properties between the carbon fibers and the matrix; however, in many cases these treatments can induce severe damage to the fiber surface, which ultimately leads to a weakening of the structure.

An attempt was made to modify the chemical, physical, and mechanical properties of carbon fibers via the catalytic growth of carbon nanofibers on their surfaces (Downs and Baker 1991, 1995). In this work the emphasis was placed on establishing conditions for the optimum growth of carbon nanofibers without inducing a corresponding degradation in the mechanical properties of the parent fibers. Using this ‘‘whiskerization’’ protocol, tests demonstrated that it was possible to achieve an improvement of over 4.5 times in the interfacial shear strength of the fibers following the deposition of a critical amount of nanofibers.

In a further investigation it was shown that the incorporation of carbon nanofibers into various epoxy resins had a dramatic impact on the curing reaction; small amounts of the filler tended to retard the process, whereas large amounts of nanofibers exerted the opposite effect (Yin *et al.* 1993). Modifications in the curing characteristics were compared with those produced from the incorporation of carbon black into the same resins, and major improvements in both the mechanical and electrical properties of the composites were realized when large amounts of carbon nanofibers were present in the systems.

The failure to create a practical storage system has prevented hydrogen from reaching the commercial forefront as a transportation fuel. The ideal hydrogen storage unit should be lightweight, compact, relatively inexpensive, safe, and reusable without the need for complex regeneration treatments. Four methods are currently being considered for hydrogen storage in commercial applications: (i) pressurized gas storage, (ii) liquefied hydrogen, (iii) selected metal hydrides, and (iv) refrigerated superactivated carbon. While the latter two approaches may offer advantages over the other technologies with regard to cost and safety aspects, they do have their own drawbacks. Metal hydrides are heavy, expensive, and release heat during the hydrogen absorption process. Active carbons are very effective adsorption agents for a variety of gases; however, interactions between the adsorbent and the solid are only of a physical nature, and as such, retention of hydrogen can only be achieved at cryogenic temperatures. It is clear that both storage and controlled release of hydrogen at moderate conditions of temperature and pressure are major issues that must be addressed if this technology is to be exploited commercially. The ideal solution would be an inexpensive, lightweight material capable of not only absorbing and retaining hydrogen at room temperature and moderate pressures, but also possessing the ability to release the gas at moderate temperatures. One type of material that fulfills these criteria is highly graphitic carbon nanofibers.

It has been demonstrated that when selected types of carbon nanofibers are pretreated under conditions whereby all adsorbed and absorbed gases are eliminated, then on subsequent exposure to hydrogen at moderate pressures, they are capable of absorbing and retaining up to 40% by weight of molecular hydrogen per gram of carbon at room temperature (Chambers *et al.* 1998b). The extraordinary hydrogen adsorption capacity exhibited by carbon nanofibers is attributed to the unique structural characteristics exhibited by these materials, which consist of an infinite number of graphite layers stacked at particular angles with respect to the fiber axis. Another important factor encountered with carbon nanofibers is that since the graphite platelets are held together by van der Waals forces and not by chemical bonds, the walls of the nanopores can readily undergo expansion to accommodate hydrogen in a multilayer fashion.

See also: Carbon Nanotubes*

Bibliography

Alstrup I 1998 A new model explaining carbon filament growth on metal catalysts. *J. Catal.* **109**, 241–51

Baker R T K 1995 A review of in-situ electron microscopy studies of metal/metal oxide–graphite interactions. *J. Adhesion* **52**, 13–40

Baker R T K, Barber M A, Harris P S, Feates F S, Waite R J 1972 Nucleation and growth of carbon deposits from the nickel catalyzed decomposition of acetylene. *J. Catal.* **26**, 51–62

Chambers A, Baker R T K 1997 Influence of chlorine on the decomposition of ethylene over iron and cobalt particles. *J. Phys. Chem. B* **101**, 1621–30

Chambers A, Nemes T, Rodriguez N M, Baker R T K 1998a Catalytic behavior of graphite nanofiber supported nickel particles: 1. Comparison with other support media. *J. Phys. Chem. B* **102**, 2251–8

Chambers A, Park C, Baker R T K, Rodriguez N M 1998b Hydrogen storage in graphite nanofibers. *J. Phys. Chem. B* **102**, 4253–6

Chambers A, Rodriguez N M, Baker R T K 1996 Influence of silver addition on the catalytic behavior of cobalt. *J. Phys. Chem.* **100**, 4229–36

Downs W B, Baker R T K 1991 Novel carbon fiber–carbon filament structures. *Carbon* **29**, 1173–9

Downs W B, Baker R T K 1995 Modification of the surface properties of carbon fibers via the catalytic growth of carbon nanofibers. *J. Mater. Res.* **10**, 625–33

Dresselhaus M S, Dresselhaus G, Sugihara K, Spain I L, Goldberg H A 1988 *Graphite Fibers and Filaments*. Springer Series in Materials Science 5, Springer, New York

Esumi K, Kimura Y, Nayada T, Meguro K, Honda H 1989 Adsorption of some solutes on heat-treated meso-carbon microbeads treated by plasma. *Carbon* **27**, 301–3

Fenelonov V B, Avdeeva L B, Goncharova O V, Okkel L G, Simonov P A, Derevyankin A Y, Likholobov V A 1995 Catalytic filamentous carbon as adsorbent and catalyst support. In: Poncelet G, et al. (eds.) *Preparation of Catalysts VI*. Elsevier, Amsterdam, pp. 825–32

Goodman D W, Kelley R D, Madey T E, Yates J T Jr. 1980 Kinetics of the hydrogenation of CO over a single crystal nickel catalyst. *J. Catal.* **63**, 226–34

Hoogenraad M S, van Leeuwarden R A G M M, van Breda Vriesman, G J B, Broersma A, van Dillen A J, Geus J W 1995 Metal catalysts supported on a novel Carbon support. In: Poncelet G, et al. (eds.) *Preparation of Catalysts VI*. Elsevier, Amsterdam, 263–71

Kim M S, Rodriguez N M, Baker R T K 1993 The interplay between sulfur adsorption and carbon deposition on cobalt catalysts. *J. Catal.* **143**, 449–63

Krishnankutty N, Rodriguez N M, Baker R T K 1996 Effect of copper on the decomposition of ethylene over an iron catalyst. *J. Catal.* **158**, 217–27

Motojima S, Kawaguchi M, Nozaki K, Iwanaga H 1991 Preparation of coiled carbon fibers by catalytic pyrolysis of acetylene and its morphology and extension characteristics. *Carbon* **29**, 379–85

Nakamura J, Hirano H, Xie M, Matsuo I, Yamada T, Tanaka K 1989 Formation of a hybrid surface of carbide and graphite layers on Ni(100) but no hybrid surface on Ni(111). *Surf. Sci.* **222**, L809–17

Nemes T, Chambers A, Baker R T K 1998 Characteristics of carbon filament formation from the interaction of cobalt-tin particles with ethylene. *J. Phys. Chem. B* **102**, 6323–30

Owens W T, Rodriguez N M, Baker R T K 1994 Effect of sulfur on the interaction of nickel with ethylene. *Catal. Today* **21**, 3–22

Park C, Baker R T K 1998 Catalytic behavior of graphite nanofiber supported nickel particles: 2. The influence of the nanofiber structure. *J. Phys. Chem.* **102**, 5168–77

Park C, Engel E, Crowe A, Gilbert T R, Rodriguez N M 1999 The use of carbon nanofibers in the removal of organic solvents from water. *Langmuir*, in press

Planeix J M, Coustel N, Coq B, Brotons V, Kumbhar P, Dutartre R, Geneste P, Bernier P, Ajayan P M 1994 Application of carbon nanotubes as supports in heterogeneous catalysis. *J. Am. Chem. Soc.* **116**, 7935–6

Reyniers M S G, Froment G F 1995 Influence of metal surface and sulfur addition on coke deposition in the thermal cracking of hydrocarbons. *Ind. Eng. Chem. Res.* **34**, 773–85

Rodriguez N M 1993 A review of catalytically grown carbon nanofibers. *J. Mater. Res.* **8**, 3233–50

Rodriguez N M, Chambers A, Baker R T K 1995 Catalytic engineering of carbon nanostructures. *Langmuir* **11**, 3862–6

Rodriguez N M, Kim M S, Baker R T K 1993 Promotional effect of carbon monoxide on the decomposition of ethylene over an iron catalyst. *J. Catal.* **144**, 93–108

Rodriguez N M, Kim M S, Baker R T K 1994 Carbon nanofibers: a unique catalyst support medium. *J. Phys. Chem.* **98**, 13108–11

Rostrup-Nielsen J R 1984 Sulfur-passivated nickel catalysts for carbon-free steam reforming of methane. *J. Catal.* **85**, 31–43

Salman F, Park C, Baker R T K 1999 Hydrogenation of crotonaldehyde over graphite nanofiber supported nickel. *Catal. Today* **53**, 385–94

Whitesides G M, Weisbecker C S 1994 Measurements of the conductivity of individual 10 nm carbon nanotubes. *Mater. Res. Soc. Symp. Proc.* **349**, 263–8

Yang R T, Chen J P 1985 Mechanism of carbon filament growth on metals. *J. Catal.* **115**, 52–60

Yin M, Koutsky J A, Barr T L, Rodriguez N M, Baker R T K, Klebanov L 1993 Characterization of carbon microfibers as a reinforcement for epoxy resins. *Chem. Mater.* **5**, 1024–31

R. T. K. Baker
Northeastern University, Boston, Massachusetts, USA

Carbon Nanotubes

Introduction

This article addresses the key issues on hand in carbon nanotube technology. It is apparent that carbon nanotubes work extremely well as field-emission sources and that the semiconducting variety of carbon nanotubes will be the ultimate ultraminiature field effect transistors that will be beneficial to a very large number of applications such as electronic devices and chemical sensors. The key challenges at present are to synthesize or separate carbon nanotubes according to their specific electronic properties so that they may be commercially viable for these envisaged applications. This article provides an overview of the rich variety of research and development in progress to attain these goals.

The discovery, nearly two decades ago, of the third ordered allotrope of carbon, referred to as buckyballs (or buckminsterfullerenes or C_{60}) by Kroto *et al.* (1985), opened the doors to the world of nanoscience and technology. It was this development that led to the discovery of other fascinating forms of nanocarbons, such as multiwalled carbon nanotubes by Iijima (1991) and, subsequently, the single-walled variant almost simultaneously by Iijima and Ichihashi (1993) and Bethune *et al.* (1993). Since the 1990s, there has been extensive publication by a large number of researchers around the world in this field, and thus it has become extremely evident that the single-walled variant of carbon nanotubes (hereafter referred to as SWNTs) is an important component of the future of materials science. SWNTs can be considered as unique materials for several reasons: they are prototype one-dimensional quantum wires that are made up of one element (carbon), that are one atom in wall thickness and tens of atoms in circumference, with every atom on the surface of the tube. The theoretical as well as experimentally measured values of the mechanical properties of SWNTs, with elastic modulus and tensile strength calculated to be near 1000 and several tens of gigapascals (GPa), respectively, are comparable to those predicted for single sheets of graphite.

Perhaps the most exciting characteristics of SWNTs are their electronic properties, as reported by Avouris (2002), which depend exclusively on their diameter and chirality (the manner in which the single sheet of hexagonally arrayed carbon atoms is wound). The electronic structure and properties of SWNTs are based on those of a single layer of graphite and its own peculiar electronic structure, and can be broadly classified as being either metallic or semiconducting, which in turn can be further classified on the basis of individual tube diameters. The bandgap of a semiconducting tube (a critical attribute to be considered for nanoelectronic applications) is inversely proportional to the tube diameter. As described in this article, research on carbon nanotubes, with particular emphasis on SWNTs, clearly indicates that these materials are going to be key players in the arena of "molecular electronics," in which the active part of the device is composed of either a single molecule or a few. However, several technical challenges need to be overcome before carbon nanotubes can be applied to electronic devices as commonly as silicon is nowadays. This article addresses some of these challenges and how the research community at the academic and industrial levels is attempting to overcome them.

1. Electronic Properties of SWNTs

The field of electronic properties of carbon nanotubes has perhaps been the most active area of research since their discovery in the early 1990s. Theoretical studies on the electronic properties of carbon nanotubes have been made using band structure calculations similar to those for planar single sheets of graphite, as described by Mintmire *et al.* (1992) and Saito *et al.* (1992). Considering the one-dimensional structural characteristics of nanotubes, the number of allowed electron states in the axial direction will be large, whereas it would be significantly limited in the circumferential direction. In these studies, using conventional indices (n, m) to represent the chirality of the individual tubes, the electronic structure of the nanotube can be calculated on a basis analogous to that of a single sheet of graphite, with the additional imposition of the necessary boundary conditions due to the limits on the nanotube circumference. In terms of their electronic properties, carbon nanotubes can therefore be classified into two major groups—those with a moderate bandgap with $n - m \neq 3p$ (where n and m are the tube indices and p is an integer) and are electrically semiconducting, and those with a small bandgap that satisfy the relation $n - m = 3p$ and are electrically metallic in nature. These findings suggest that armchair (n, n) tubes would be metallic under all circumstances, and that zigzag (n, 0) and chiral (n, m) tubes would be metallic in one-third of the cases and semiconducting in the other two-thirds. It is also important to note that several other theoretical studies have demonstrated that "defect" structures on the tubes would lead to metal–semiconductor, semiconductor–semiconductor, and metal–metal junctions at the defect sites, which might themselves behave as nanoscale device elements, as demonstrated by Chico *et al.* (1996). Topological defects and atomic impurities in the tubes alter the local electronic density-of-states near the Fermi level, leading to localized variations in the dominant conductivity associated with the individual tube.

Recently, several research groups have attempted to measure the electrical properties of carbon nanotubes experimentally. The development of scanning tunneling microscopy (STM) and spectroscopy-based techniques has been particularly helpful in these experimental developments. The initial step of the experimental approach involves the dispersing of dilute suspensions of nanotubes onto patterned electrical contacts with the intention of contacting single nanotubes to two adjacent electrodes for electrical transport measurements. Considering that electrons in π-state control the electrical transport in nanotubes, the low-bias resistance of the nanotube for ballistic transport would be $h/4e^2$, which is ~ 6 kΩ. However, the two-point resistance of single-nanotube samples draped over platinum electrodes (or gold electrodes in other cases), which is shown in Fig. 1, was initially measured to be closer to 1 MΩ at room temperature, as reported by Tans *et al.* (1997). The key lesson from this initial experiment was that the contact resistance between the carbon nanotube and the metal electrode was a major factor in determining resistivity, and this was an issue that needed further consideration in future experiments. Also, the localized changing of the bond structure of the carbon atoms from sp^2-like to sp^3-like, when the nanotube was "draped" over the metal electrodes, led to significant resistivity enhancement; this was another serious technical issue to be addressed. But it was apparent from the initial experimental analysis of electronic properties that carbon nanotubes behaved as coherent quantum wires and that the structure (chirality) was indeed a major factor in determining whether individual tubes were metallic or semiconducting (Wildöer *et al.* 1998). Recent experiments have also demonstrated that the conductivity of individual SWNTs depends significantly on adsorbed gases such as oxygen, making pure SWNTs extremely sensitive detectors for oxygen (Collins *et al.* 2000). These elegant experiments clearly confirmed the theoretical work on the electronic properties of carbon nanotubes, as did other experiments that confirmed that defect structures led to substantial enhancement in the resistivity of carbon nanotubes, as had been previously shown by Dai *et al.* (1996).

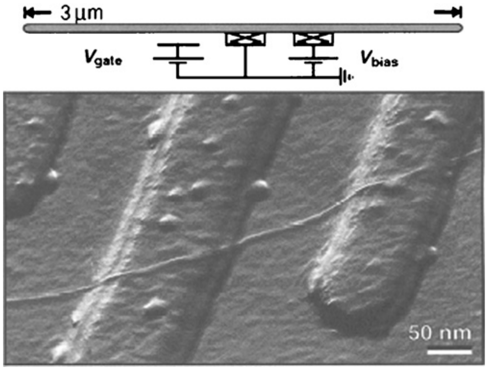


Figure 1
Atomic force microscope (AFM) tapping-mode image of a single SWNT draped over two 15 nm-thick Pt-electrodes on an Si/SiO$_2$ substrate and a corresponding circuit diagram. Reprinted by permission from Tans S J, Devoret M H, Dai H, Thess A, Smalley R E, Geerligs L J, Dekker C 1997 Individual single-wall carbon nanotubes as quantum wires. *Nature* **386**, 474–7 copyright 1997 Macmillan Publishers Ltd.

2. Dispersion and Separation of SWNT

Probably, the major concern for rendering SWNTs commercially viable for electronic devices is to be able to separate one tube from the next and to be able to classify them according to their electronic properties, such as whether they are metallic ballistic conductors or semiconductors. It would also be desirable to be able to separate the semiconducting nanotubes according to their respective chirality, which would be particularly important for nanoelectronic applications. Currently, SWNTs are synthesized by one of three different techniques—laser ablation, gas-phase catalytic growth from carbon monoxide (CO), or by chemical vapor deposition (CVD) from other hydrocarbon precursors, typically using nanometer-scale transition metal catalysts mostly supported on oxides such as silica. The laser ablation and gas-phase catalytic growth from CO (referred to as HiPCO), techniques have been described in detail by Thess *et al.* (1996) and Nikolaev *et al.* (1999). Being a continuous flow process rather than a batch process, the HiPCO method offered the advantage of being able to be scaled up to produce SWNTs in kilogram quantities per day.

The processes described above, however, produced materials fraught with impurities, such as large amounts of disordered or glassy carbon in the case of the laser-ablation process or carbon-encapsulated catalyst iron particles adhering to the walls of the tubes in the HiPCO process. In either case, the tubes themselves were found to be a mixture of metallic and semiconducting types. In order to render SWNTs potentially useful for high-performance electronic devices such as field effect transistors or sensors, it is necessary to be able to obtain tubes that are mostly, or preferably, ''only'' semiconducting in electronic structure. The options available are processes that would enable high degrees of separation of the metallic tubes from the semiconducting ones. The ultimate ''holy grail'' would be a process or processes that selectively grow only semiconducting tubes (or preferably, semiconducting tubes with one chirality or a narrow set of chiralities) or only metallic tubes as the need may be. Analytical tools such as spectrofluorometric analysis in combination with resonance Raman scattering have already been demonstrated to map the chiralities of individual classes of SWNTs in a product, as illustrated by Bachilo *et al.* (2002) and reproduced in the on-line edition of the Encyclopedia of Materials: Science and Engineering, doi: 10.1016/B0-08-043152-6/00179-0.

3. Functionalization of Carbon Nanotubes to Enhance Selection and Separation

As-synthesized carbon nanotubes are highly polarizable, smooth-sided materials with attractive interactions of $\sim 0.5\,\mathrm{eV\,nm^{-1}}$ for tube-to-tube contact. At the dimensions of SWNTs, these cohesive forces are sizeable and cause the tubes to agglomerate into bundles. Inherently, these bundles are extremely difficult to separate into individual tubes for further manipulation or any potential uses of the SWNTs in other devices. Several types of functionalization schemes have been studied to potentially overcome these problems. Noncovalent functionalization, as for instance by the use of surfactants such as sodium dodecyl sulfate (SDS), leads to the identification of tube chiralities by fluorescence spectroscopy (O'Connell *et al.* 2002). Strategies for covalent functionalization of carbon nanotubes include defect creation or attachment of carboxylic acid groups to end-cap sites in order to provide local anchoring sites for the functional groups, or direct covalent side-wall functionalization. The methodologies and resultant added groups attached to the walls of the carbon nanotubes have been reported by a number of researchers. The addition of preformed arene diazonium salts to micelle-coated nanotubes produces high levels of functionalization of aryl groups (estimated by Raman spectroscopy and thermogravimetric analyses to be as high as 10%) on the tube walls, as reported by Bahr *et al.* (2001). Functionalization of the tube walls at such levels causes the tubes to disperse well and stay dispersed in organic solvents. Figure 2 illustrates a transmission electron microscopic image of a single nanotube, its walls appearing rougher than typical nanotubes due to functionalization by aryl moieties. Other functionalization schemes that have been investigated are fluorination by Mickelson *et al.* (1998) and radical chemistry by Ying *et al.* (2003). A relatively lighter functionalization scheme was discussed by Georgakilas *et al.* (2002) involving a 1,3-dipolar cycloaddition of azomethyne ylide generated by the reaction of amino acid and aldehyde.

Recently, there have also been a number of both covalent and noncovalent functionalization chemistries that may be used to preferentially select metallic tubes over semiconducting tubes. The most striking one among these studies is the one reported by Zheng

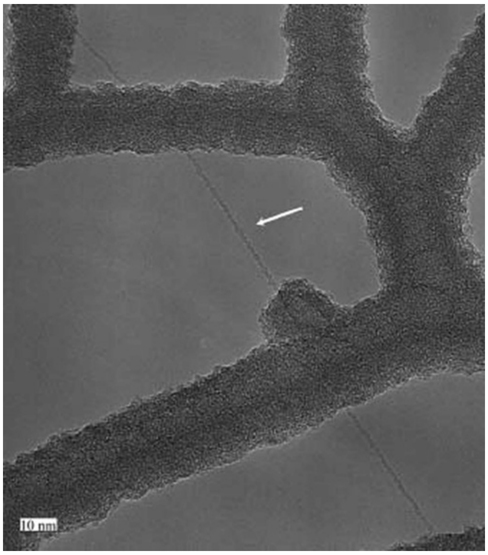


Figure 2
High-resolution TEM image of a single SWNT functionalized by 4-*tert*-butylbenzene. Reprinted with permission from Dyke C A, Tour J M 2004 Covalent functionalization of single-walled carbon nanotubes for materials applications. *J. Phys. Chem. A* **108**(51), 11151–9. Copyright 2004 American Chemical Society.

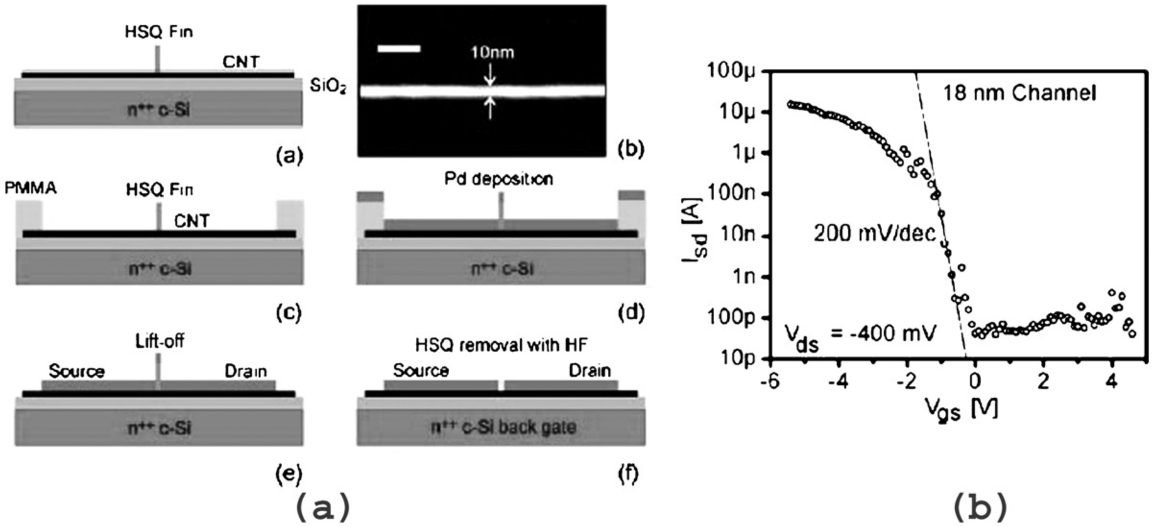

HSQ Fin
CNT
SiO2
n++ c-Si
(a)
10nm
(b)
PMMA
HSQ Fin
CNT
n++ c-Si
(c)
Pd deposition
n++ c-Si
(d)
Lift-off
Source
Drain
n++ c-Si
(e)
HSQ removal with HF
Source
Drain
n++ c-Si back gate
(f)
(a)
100µ
10µ
1µ
100n
10n
1n
100p
10p
Isd [A]
18 nm Channel
200 mV/dec
Vds = -400 mV
-6
-4
-2
0
2
4
Vgs [V]
(b)

Figure 3
(a) Process-flow schematic for the creation of short-channel field effect transistors using SWNT and (b) gate-voltage dependence of the source–drain current in an ~18 nm channel field effect transistor for a drain–source voltage of −0.4 V. Reprinted with permission from Seidel R V, Graham A P, Kretz J, Rajasekharan B, Duesberg G S, Liebau M, Unger E, Kreupl F, Hoenlein W 2005 Sub-20 nm short channel carbon nanotube transistors. *Nano Lett*. **5**(1), 147–50. Copyright 2005 American Chemical Society.

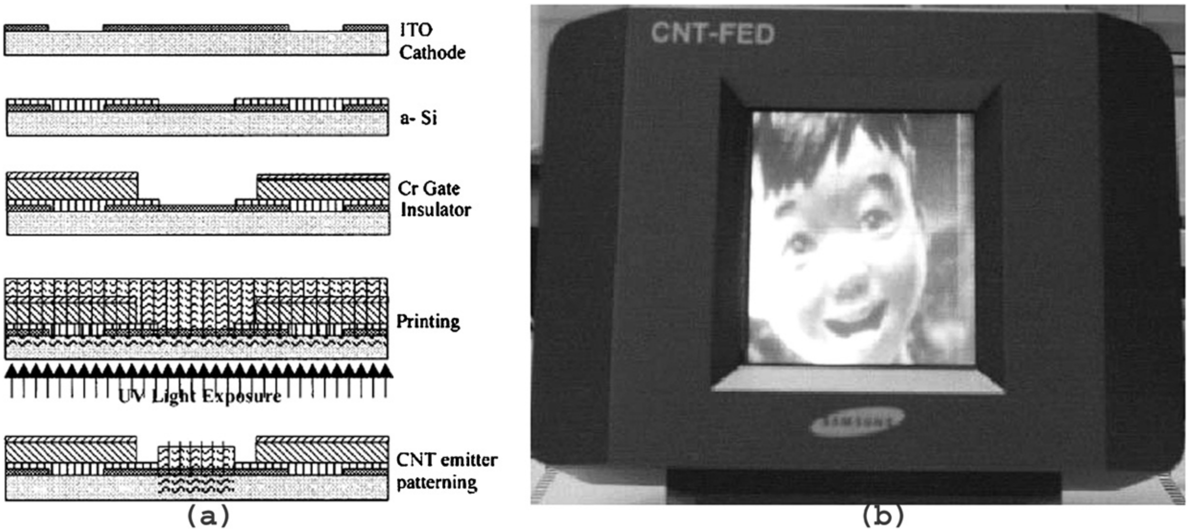

ITO
Cathode
a- Si
Cr Gate
Insulator
Printing
UV Light Exposure
CNT emitter
patterning
(a)
CNT-FED
SAMSUNG
(b)

Figure 4
(a) Schematic of the field-emitting display with a gated structure developed at Samsung Corporation with SWNT as the emission source. Reprinted with permission from Deuk-Seok Chung, Park S H, Lee H W, Choi J H, Cha S N, Kim J W, Jang J E, Min K W, Cho S H, Yoon M J, Lee J S, Lee C K, Yoo J K, Jong-Min Kim, Jung J E, *et al.* 2002 Carbon nanotube electron emitters with a gated structure using backside exposure processes. *Appl. Phys. Lett*. **80**(21), 4045–7 (2002). Copyright 2002, American Institute of Physics. (b) Color video running image with a 5-inch diagonal active display area.

et al. (2003), showing that single-stranded DNA molecules with specific sequences can self-assemble onto the highly ordered surfaces of SWNTs, which allows not only for the separation of metallic from semiconducting tubes but also for separation of semiconducting tubes based on individual diameters.

4. Device Fabrication

Of the various existing and potential applications involving carbon nanotubes, none appears to evoke so much interest or excitement as the use of single nanotubes (particularly single semiconducting SWNTs), as field effect transistors (FETs) at room temperature. What is driving this particular application/technology? Conventional semiconductor lithographic techniques are reaching their theoretical, practical, and possibly economic limits. The demand /value for continuing shrinkage of scale will probably keep growing even after conventional structures have reached these limits. From a technological standpoint, *n*-type FETs, where transport through the semiconducting SWNTs is dominated by electrons and not by holes, allow the fabrication of nanotube-based complementary logic devices and circuits. This can be achieved by converting the predominantly *p*-type SWNTs to *n*-type SWNTs by either alkali-metal doping or by annealing the device either in vacuum or in an inert atmosphere, which results in oxygen removal that subsequently alters the line-up of the valence bands at the electrode–nanotube junction (Derycke *et al.* 2002). Other efforts to develop this technology include lowering the contact resistance of the nanotube–electrode junction by "end-bonding" the nanotube to metals such as titanium (Martel *et al.* 2001), and altering the configuration of the gate electrodes to improve the overall electrical characteristics (Wind *et al.* 2002). Most recently, researchers in Infineon Technologies AG have successfully demonstrated the fabrication and operation of a sub-20 nm FET based on SWNTs (Seidel *et al.* 2005) using electron beam lithography with select e-beam resists. This device performs as well as much larger state-of-the-art silicon-based devices in many aspects with high on/off current ratios. A schematic of the process flow for this device and its current–voltage performance is illustrated in Fig. 3.

5. Advances in Field-emission Display Technology

The unique combination of the physical structure (high aspect ratio and nanometer-scale radius of curvature of the tip), mechanical properties such as stiffness and strength, and the chemical stability of the single graphite sheet that constitutes its surface makes the carbon nanotube a consummate field-emitting device. Several research groups have demonstrated that thin ropes of single-walled nanotubes are superb field-emitting sources that show high emission properties (lower turn-on voltage and higher emissivity). The heart of such devices is the cathode, where the nanotubes serve as emission sources following appropriate fabrication techniques. In prototype demonstrations using a 4.5-inch diode-based flat panel display with ropes of single-walled tubes serving as emitting sources, researchers in the field-emitting display program at Samsung have demonstrated that the device can be turned on at low voltages (less than $1\,\mathrm{V\,\mu m^{-1}}$), and the brightness measured on the green phosphor (ZnS:Cu, Al) of the screen was $\sim$1800 candelas per square meter at $4\,\mathrm{V\,\mu m^{-1}}$ (Choi *et al.* 1999). This development appears to be well on its way to commercialization with the recent demonstration of SWNT-based field emitters with gated structures, which allow the addressing of individual pixels and the subsequent development of video running image displays. The schematic of the field-emitting display

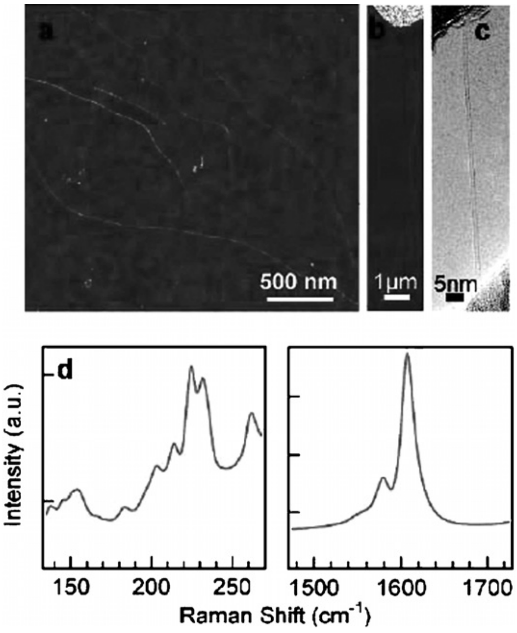


Figure 5
Analytical data pertaining to SWNT grown by PECVD: (a and b) AFM images of SWNT grown from low-density ferritin deposition on SiO_2 and from an iron-film island, respectively, (c) TEM of an as-grown SWNT, and (d) Raman spectra for the radial breathing modes of SWNT and *G*-band of SWNT vibration with clear peaks at $1578\,\mathrm{cm^{-1}}$ and $1608\,\mathrm{cm^{-1}}$. Reprinted with permission from Li Y, Mann D, Rolandi M, Kim W, Ural A, Hung S, Javey A, Cao J, Wang D, Yenilmez E, Wang Q, Gibbons J F, Nishi Y, Dai H 2004 Preferential growth of semiconducting single-walled carbon nanotubes by a plasma enhanced CVD method. *Nano Lett.* **4**(2), 317–21. Copyright 2004 American Chemical Society.

with SWNTs as emission sources with a gated structure and the flat panel display video image obtained from it are shown in Fig. 4. However, in the huge market for high-definition televisions, currently in excess of US $60 billion annually, carbon nanotube-based field-emission display televisions would have to compete against other existing and rapidly growing technologies such as plasma and liquid crystal displays. Several factors such as cost of manufacture, stability, power usage, and viewer-friendliness will ultimately determine the future of carbon nanotube-based field-emission technology in the long run.

6. Latest Advances in Synthesis Techniques

While it may appear that the research community at large is working on finding applications for carbon nanotubes, efforts to synthesize them with the desired structure or without impurities are still underway. Most recently, Li *et al.* (2004) have demonstrated a relatively low temperature (600 °C) plasma-enhanced CVD technique to produce SWNTs that are nearly 90% semiconductors. Analysis of this product by microscopy and Raman spectroscopy reveals that the nanotubes are of high quality, as illustrated in Fig. 5.

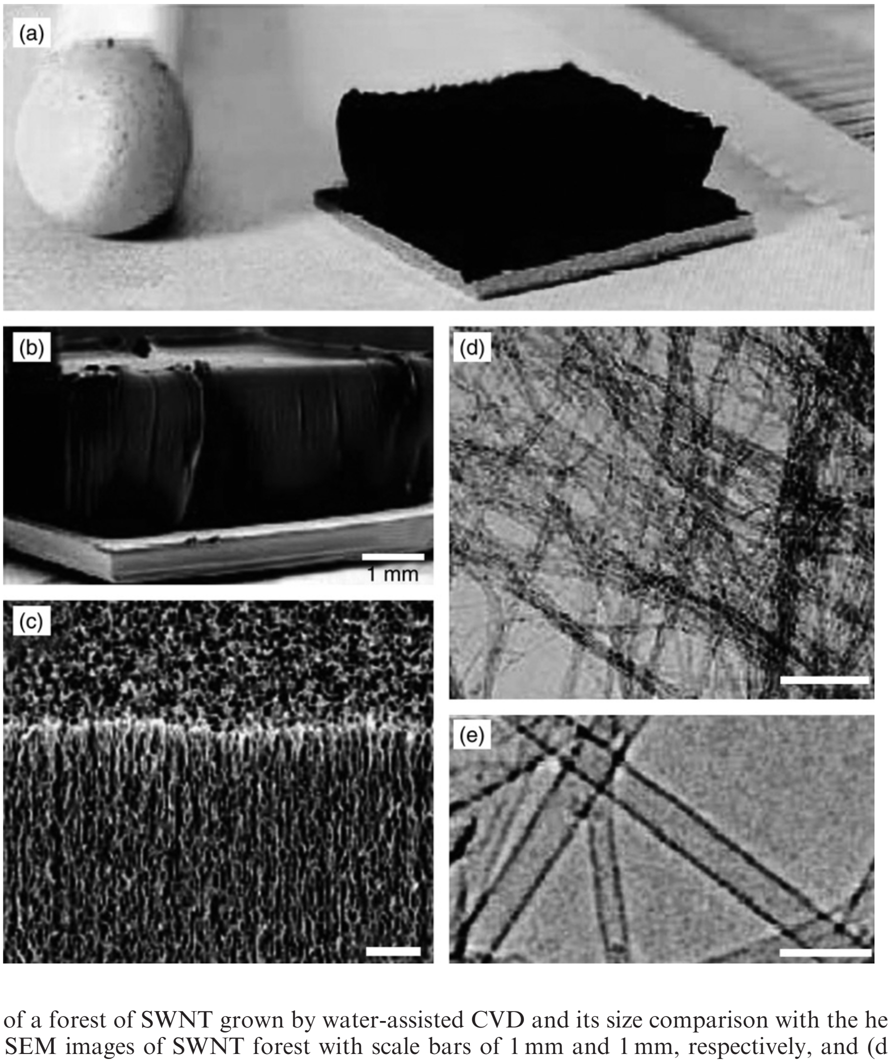


Figure 6
(a) Photograph of a forest of SWNT grown by water-assisted CVD and its size comparison with the head of a matchstick, (b and c) SEM images of SWNT forest with scale bars of 1 mm and 1 mm, respectively, and (d and e) TEM images of the SWNT with scale bars of 100 nm and 5 nm, respectively. Reprinted with permission from Hata K, Futaba D N, Mizuno K, Namai T, Yumura M, Iijima S 2004 Water-assisted highly efficient synthesis of impurity-free single-walled carbon nanotubes. *Science* **306**, 1362–49. Copyright 2004 AAAS. Permission from AAAS is required for all other uses.

CVD, with the assistance of water, was also used recently to demonstrate that thick and aligned "forests" of SWNT can be grown where the catalysts can be easily separated from the tubes to provide carbon purities well in excess of 99% (Hata *et al.* 2004)—images of this new material are shown in Fig. 6. Although this latest offering has not yet been tested for any particular properties, it is easily conceivable that it could find use in a number of commercial applications.

Bibliography

Avouris P 2002 Molecular electronics with carbon nanotubes. *Acc. Chem. Res.* **35**, 1026–34

Bachilo S M, Strano M S, Kittrell C, Hauge R H, Smalley R E, Weisman R B 2002 Structure-assigned optical spectra of single-walled carbon nanotubes. *Science* **298**, 2361–6

Bahr J L, Yang J, Kosynkin D V, Bronikowski M J, Smalley R E, Tour J M 2001 Functionalization of carbon nanotubes by electrochemical reduction of aryl diazonium salts: a bucky paper electrode. *J. Am. Chem. Soc.* **123** (27), 6536–42

Bethune D S, Kiang C H, de Vries M S, Gorman G, Savoy R, Vazquez J, Beyers R 1993 Cobalt-catalyzed growth of carbon nanotubes with single-atomic-layer walls. *Nature* **363**, 605–7

Chico L, Crespi V H, Benedict L X, Louie S G, Cohen M L 1996 Pure carbon nanoscale devices: nanotube heterojunctions. *Phys. Rev. Lett.* **76**, 971–4

Choi W B, Chung D S, Kang J H, Kim H Y, Jin Y W, Han I T, Lee Y H, Jung J E, Lee N S, Park G S, Kim J M 1999 Fully sealed, high-brightness carbon nanotube field emission display. *Appl. Phys. Lett.* **75** (20), 3129–31

Collins P G, Bradley K, Ishigami M, Zettl A 2000 Extreme oxygen sensitivity of electronic properties of carbon nanotubes. *Science* **287**, 1801–4

Dai H, Wong E W, Lieber C M 1996 Probing electrical transport in nanomaterials: conductivity of individual carbon nanotubes. *Science* **272**, 523–6

Derycke V, Martel R, Appenzeller J, Avouris P 2002 Controlling doping and carrier injection in carbon nanotube transistors. *Appl. Phys. Lett.* **80** (15), 2773–5

Deuk-Seok Chung, Park S H, Lee H W, Choi J H, Cha S N, Kim J W, Jang J E, Min K W, Cho S H, Yoon M J, Lee J S, Lee C K, Yoo J H, Jong-Min Kim, Jung J E, *et al.* 2002 Carbon nanotube electron emitters with a gated structure using backside exposure processes. *Appl. Phys. Lett.* **80** (21), 4045–7

Dresselhaus M S, Dresselhaus G, Eklund P C 1995 *Science of Fullerenes and Carbon Nanotubes*. Academic Press, San Diego, CA

Dresselhaus M S, Dresselhaus G, Avouris P (eds.) 2001 *Carbon Nanotubes*: synthesis, structure, properties and applications. Springer, Berlin

Dyke C A, Tour J M 2004 Covalent functionalization of single-walled carbon nanotubes for materials applications. *J. Phys. Chem. A* **108** (51), 11151–9

Georgakilas V, Kordatos K, Prato M, Guldi D M, Holzinger M, Hirsch A 2002 Organic functionalization of carbon nanotubes. *J. Am. Chem. Soc.* **124** (5), 760–1

Hata K, Futaba D N, Mizuno K, Namai T, Yumura M, Iijima S 2004 Water-assisted highly efficient synthesis of impurity-free single-walled carbon nanotubes. *Science* **306**, 1362–4

Iijima S 1991 Helical microtubules of graphitic carbon. *Nature* **354**, 56–8

Iijima S, Ichihashi T 1993 Single-shell carbon nanotubes of 1-nm diameter. *Nature* **363**, 603–5

Kroto H W, Heath J R, O'Brien S C, Curl R F, Smalley R E 1985 C_{60}: buckminsterfullerene. *Nature* **318**, 162–3

Li Y, Mann D, Rolandi M, Kim W, Ural A, Hung S, Javey A, Cao J, Wang D, Yenilmez E, Wang Q, Gibbons J F, Nishi Y, Dai H 2004 Preferential growth of semiconducting single-walled carbon nanotubes by a plasma enhanced CVD method. *Nano Lett.* **4** (2), 317–21

Martel R, Derycke V, Lavoie C, Appenzeller J, Chan K K, Tersoff J, Avouris P 2001 Ambipolar electrical transport in semiconducting single-wall carbon nanotubes. *Phys. Rev. Lett.* **87** (25), 256805 -1–4.

Mickelson E T, Huffman C B, Rinzler A G, Smalley R E, Hauge R H, Margrave J L 1998 Fluorination of single-wall carbon nanotubes. *Chem. Phys. Lett.* **296**, 188–94

Mintmire J W, Dunlap B I, White C T 1992 Are fullerene tubules metallic? *Phys. Rev. Lett.* **68**, 631–4

Nikolaev P, Bronikowski M J, Bradley R K, Rohmund F, Colbert D T, Smith K A, Smalley R E 1999 Gas-phase catalytic growth of single-walled carbon nanotubes from carbon monoxide. *Chem. Phys. Lett.* **313**, 91–7

O'Connell M J, Bachilo S M, Huffman C B, Moore V C, Strano M S, Haroz E H, Rialon K L, Boul P J, Noon W H, Kittrell C, Ma J, Hauge R H, Weisman R B, Smalley R E 2002 Band gap fluorescence from individual single-walled carbon nanotubes. *Science* **297**, 593–6

Saito R, Fujita M, Dresselhaus G, Dresselhaus M S 1992 Electronic structure of chiral graphene tubules. *Appl. Phys. Lett.* **60**, 2204–6

Seidel R V, Graham A P, Kretz J, Rajasekharan B, Duesberg G S, Liebau M, Unger E, Kreupl F, Hoenlein W 2005 Sub-20 nm short channel carbon nanotube transistors. *Nano Lett.* **5** (1), 147–50

Tans S J, Devoret M H, Dai H, Thess A, Smalley R E, Geerligs L J, Dekker C 1997 Individual single-wall carbon nanotubes as quantum wires. *Nature* **386**, 474–7

Thess A, Lee R, Nikolaev P, Dai H, Petit P, Robert J, Xu C, Lee Y H, Kim S G, Rinzler A G, Colbert D T, Scuseria G E, Tomanek D, Fischer J E, Smalley R E 1996 Crystalline ropes of metallic carbon nanotubes. *Science* **273**, 483–7

Wildöer J W G, Venema L C, Rinzler A G, Smalley R E, Dekker C 1998 Electronic structure of atomically resolved carbon nanotubes. *Nature* **391**, 59–64

Wind S J, Appenzeller J, Martel R, Derycke V, Avouris P 2002 Vertical scaling of carbon nanotube field-effect transistors using top gate electrodes. *Appl. Phys. Lett.* **80** (20), 3817–9

Ying Y, Saini R K, Liang F, Sadana A K, Billups W E 2003 Functionalization of carbon nanotubes by free radicals. *Org. Lett.* **5** (9), 1471–3

Zheng M, Jagota A, Strano M S, Santos A P, Barone P, Chou S G, Diner B A, Dresselhaus M S, Mclean R S, Onoa G B, Samsonidze G G, Semke E D, Usrey M, Walls D J 2003 Structure-based carbon nanotube sorting by sequence-dependent DNA assembly. *Science* **302**, 1545–8

S. Subramoney
DuPont Company, Wilmington, Delaware, USA

Carbon–Carbon Composites

For many applications the ideal structural material would have high strength and stiffness and low density. For high-temperature service, the material should also retain its mechanical properties at elevated temperature, be resistant to thermal shock, and be resistant to creep. Perhaps better than any material, carbon–carbon fulfills these requirements. Carbon–carbon is the most highly developed ceramic composite and consists of carbon fiber reinforcement of a carbon matrix as shown in Fig. 1. The fibers and matrix are evident in the scanning electron micrograph, and conical regions of matrix are seen in the optical micrographs.

Carbon–carbon is used as the friction material in aircraft brakes, nose cones of ballistic missiles, leading edges in high-performance aerospace vehicles such as the US Space Shuttle, and as turbine rotors and other high-temperature engine and rocket components. The range of applications has been broadened to include thermal management in electronic components and heat pipes where the very high thermal conductivity of certain carbon–carbon composites is advantageous. This article provides a brief history of carbon–carbon followed by a description

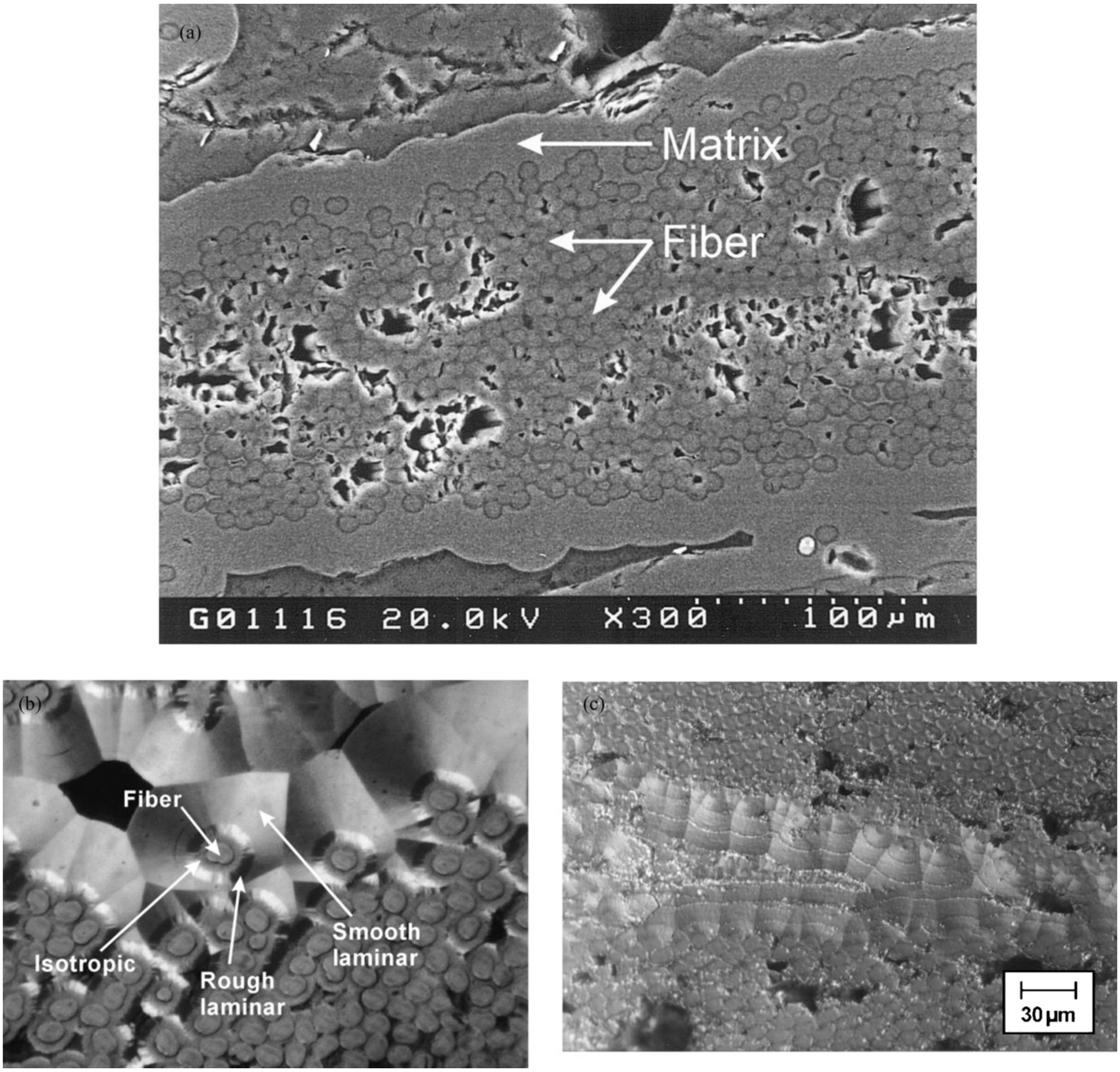


Figure 1
(a) Scanning electron, (b) optical polarized light, and (c) optical micrographs of carbon–carbon composite.

of fabrication processes. The influence of processing on the microstructure and properties of carbon–carbon are also briefly presented. Applications and cost of carbon–carbon are also reviewed. The reader is referred to several publications on carbon–carbon (Buckley and Edie 1993, Chung 1994, Fitzer and Manocha 1998, Savage 1993, Thomas 1993).

1. History of Carbon–Carbon

Graphite, the form of carbon having a hexagonal crystal structure, is an attractive material for applications at elevated temperatures in an inert atmosphere because of its high sublimation temperature, its retention of strength with increasing temperature, and its chemical inertness. However, the strength and fracture toughness of graphite are less than that desired for many structural applications. Reinforcement with strong carbon fibers, which became available in the late 1950s, improves both strength and toughness significantly.

Carbon composites were initially developed for the US Air Force space plane project and NASA's Apollo projects and received increased emphasis as a result of the US Space Shuttle program (Buckley and Edie 1993). Following this success, extensive work was performed on carbon–carbon both in the USA and Europe in the early 1970s, mainly for military applications (Warren and Williams 1972, Kotlensky 1971, 1973, Pierson 1967, Pierson *et al.* 1968, Granoff *et al.* 1973, Theis 1972, McAllister and Teverna 1972, Fitzer *et al.* 1978, Girard 1978). The original carbon composites used low-modulus rayon precursor fibers in the form of woven fabrics. The matrix was derived from the pyrolysis of phenolic and furan resins or was deposited via the chemical vapor infiltration (CVI) process which is described later. Good reviews are those of Kotlensky (1973) and Lackey and Starr (1990). Since the 1970s major developments have taken place in the production of high-modulus polyacrylonitrite (PAN) and pitch-derived fibers for use as reinforcements (Buckley and Edie 1993, Chung 1994, Fitzer and Manocha 1998).

2. Fabrication of Carbon–Carbon

Carbon–carbon composites are fabricated by starting with an array of carbon fibers referred to as the preform. The fibers typically fill about half of the preform volume. The remaining space is void until the carbon matrix is formed by either a liquid impregnation or gaseous infiltration process. The matrix holds the fiber array together and transfers mechanical loads experienced in service to the fibers.

2.1 Fiber and Preform Types

Many different types of carbon fibers are commercially available for use as the reinforcement constituent in carbon–carbons (Buckley and Edie 1993, Chung 1994). Mechanical properties such as tensile strength and physical properties such as thermal and electrical conductivity vary extensively from one fiber type to another. For example, PAN- and pitch-derived fibers are much stronger and have a higher elastic modulus than rayon-derived fiber. Even for a given fiber type, the properties and cost are strongly dependent on the highest temperature the fibers experience during processing. Higher temperatures encourage a greater percentage of the graphite-like crystals comprising the fiber to orient with their crystallographic *c*-axis perpendicular to the long axis of the fiber. Since the crystals are highly anisotropic, the properties of the fiber vary depending on the processing conditions. Increased preferred orientation of the crystals typically increases the elastic modulus as well as increases the anisotropy of the fiber. Broad ranges for tensile strength, elastic modulus, and fiber cost are, respectively, 1–5.7 GPa, 100–800 GPa, and US$2–50 per kilogram (Chung 1994). Fiber diameters are typically 7–10 μm but much smaller discontinuous fibers called "whiskers" may have a diameter less than 1 μm. For additional details on carbon fibers and whiskers see *Carbon Fibers**. Carbon whiskers produced by Applied Sciences, Inc. of Cedarville, Ohio have exceptional crystallinity as shown in Fig. 2. These whiskers have very high thermal conductivities, approaching that of diamond (Ting *et al.* 1995).

Preforms are typically comprised of carbon fibers in either continuous or chopped form. The fibers are

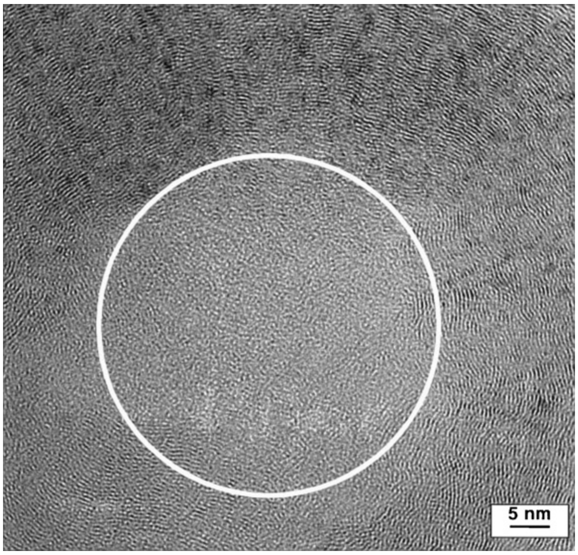


Figure 2
Highly crystalline, vapor-grown carbon whisker with basal planes oriented parallel to the surface of a central, less crystalline region.

often woven into cloth and preforms are readily fabricated by cutting and stacking the cloth layers. Adjacent layers are typically rotated 30° or 45° from one another so that fibers are oriented in several directions, although only 0–90° orientations are often used. Such two-dimensional preforms yield composites having high strength, elastic modulus, and conductivity in the plane of the component that is parallel to the cloth layers. Properties perpendicular to the lay-up plane are much inferior. Three-dimensional weaving and other textile processes can be used to orient some fibers in the transverse direction if the application requires more nearly isotropic properties.

Preforms are also regularly made by felting and slurry molding processes. The fibers or whiskers in such preforms do not all lie within one plane and thus the preforms are not as anisotropic as two-dimensional preforms. Previously the preforms for brakes were made by cloth lay-up, but cost considerations and the desire for improved properties in the transverse direction have led to the use of noncloth preforms. The latter typically have lower fiber contents than cloth lay-ups which may be as high as 60 vol.%. Several good descriptions of preform types are available (Buckley and Edie 1993, Fitzer and Manocha 1998).

2.2 *Matrix via Liquid Impregnation*

Carbon–carbon composites are fabricated using either the liquid impregnation method or CVI. In the former, the preform is impregnated with a polymer resin or coal tar or petroleum pitch which is then carbonized, i.e., heated in an oxygen-free atmosphere to cause volatilization of all elements (hydrogen, nitrogen, etc.) except carbon. Evacuation and pressurization may be used to assist flow of the liquid into the preform. Carbonization may also be performed at high pressure, such as in an autoclave. A subsequent higher temperature process step, called graphitization, which is conducted in the range 1500–3000 °C can be used to cause the carbon matrix to approach the crystal structure of graphite. A process diagram showing the steps involved is shown in Fig. 3. The resin and pitch shrink during carbonization, necessitating numerous cycles of impregnation and carbonization in order to obtain dense, i.e., low-porosity, composites. Sometimes components are initially densified using the impregnation process with CVI being used later to further reduce matrix porosity. A carbon matrix derived from pitch is more easily graphitized than when resin is used and thus pitch-derived composites are used for certain high-temperature and thermal management applications where thermal stability and/or high thermal conductivity are important. The CVI process can also yield a matrix that is readily graphitized. Carbon–carbon made by the CVI process has better structural properties than material made by multiple impregnation–carbonization cycles since shrinkage does not occur, and is thus often preferred for critical applications (Kotlensky 1971).

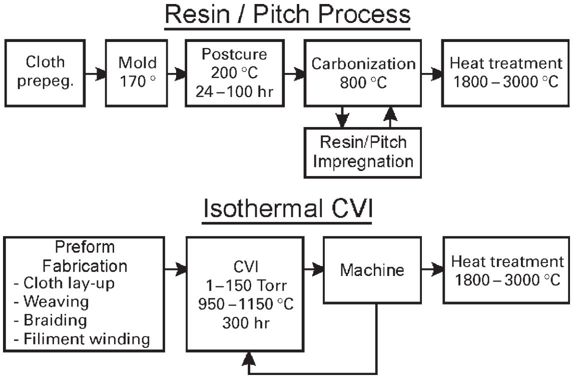


Figure 3
Processes for fabrication of carbon–carbon.

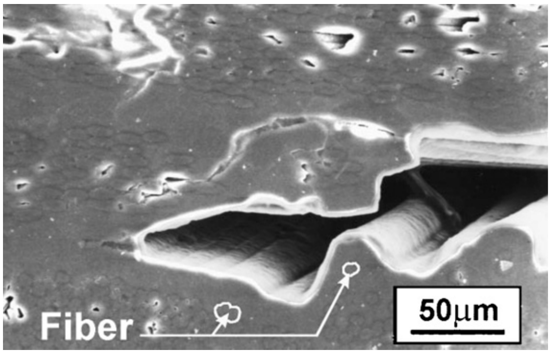


Figure 4
Micrograph of carbon–carbon prepared by CVI.

2.3 *Matrix via CVI*

The CVI process for fabricating carbon–carbon begins by placing the carbon fiber preform in a furnace (Lackey and Starr 1990). Gaseous chemical vapor deposition (CVD) reagents that are flowing inside the furnace penetrate the pores of the preform and deposit a coating of solid carbon onto the fibers. Methane, propane, or other hydrocarbons, including mixtures, may be the active reagent. Diluent gases may be hydrogen, nitrogen, or argon. As this deposition process continues, the fibers get larger and consequently the spaces between the fibers become smaller. Eventually, the fiber coatings interlock and the fibers are held together by the coating. The carbon coating is the matrix which, along with the fibers, constitute the composite. The micrograph in Fig. 4

gives an insight into the process. The contour of fibers which have been coated with carbon are visible on the surface of the pore. The original cross-sections of several fibers are outlined in white.

A brief history of the CVI process is available (Lackey and Starr 1990). The CVI process for the deposition of carbon was first used to increase the density of graphite. Similar CVI processing is used to fabricate the highest-quality carbon–carbon. There are eight variants of the CVI process, depending on whether: (i) the reagents diffuse into or are forced to flow through the preform, (ii) the preform is isothermal or exposed to a temperature gradient, and (iii) the gas pressure is uniform or pulsed by cyclic evacuation and pressurization. In addition, the preform may be heated by one of several means including resistance, radiation, and induction.

Of the eight basic processes, primarily three have been used for fabrication of carbon–carbon. Industrially, large quantities of carbon–carbon for friction applications are made using the diffusion-isothermal process without pulsing. This process is often referred to as "isothermal CVI". Industrially, several thousand brake parts are placed in a furnace having a diameter of the order of 3 m. Hydrocarbon gases continually flow through the furnace which operates at temperatures in the range 950–1150 °C and pressures of 0.1–20 kPa (Kotlensky 1973). Typical values are 1100 °C and 0.6–6 kPa. Exact processing details, including processing time, are considered proprietary.

With the isothermal CVI process, deposition occurs preferentially onto fibers located at or near the external surface of the preform. As a result, the density of the completed component is less at the center of the preform. In order to minimize the density (porosity) gradient through the thickness of the component, the process variables are selected to retard the kinetics of carbon deposition relative to the diffusion of reagent. Increased uniformity is achieved by the use of lower temperatures, pressures, and reagent concentration even though processing times are lengthened. The infiltration time can be as long as three weeks. Since deposition occurs preferentially at the surface, a dense skin forms which hinders diffusion of reagents into the component. Typically, the infiltration process is interrupted after one and two weeks to permit removal of the dense surface region by machining. Alternatively, some manufacturers choose to adjust the process variables to further reduce the kinetics of the carbon deposition reaction. This increases the time required to complete infiltration but may permit the elimination of one of the machining steps. When the infiltration process is completed, the outer denser region of the component may also be machined away in order to produce a material having minimal variation in density through the thickness of the component. The final density of carbon–carbon used in aircraft brakes is in the range 1.70–1.85 g cm^{-3} (Tatarzycki and Webb 1992).

Another type of CVI process which has been used for many years to produce carbon–carbon nose cones for ballistic missiles also uses diffusion of the reagent, but a steep thermal gradient is present through the thickness of the preform (Kotlensky 1971). One surface of a carbon fiber preform, for example the inside surface of a nose cone as shown in Fig. 5, is in contact with a graphite susceptor which is heated inductively. Hydrocarbon reagents that are flowing through the reaction vessel diffuse through the cooler outer surface of the preform toward the hotter inner surface. When the reagents become suitably heated, they decompose, forming a carbon deposit on the fibers. This densifies the inner portion of the cone and increases its thermal conductivity so that the temperature gradient through the thickness of the cone slowly changes with time. Gradually, over a period of tens of hours, the deposition front moves toward the outer surface of the preform (Theis 1972, Kotlensky 1973). This inside-out deposition process prevents the formation of a dense surface skin and parts can be made more rapidly and without interrupting the process for machining.

A third CVI process uses forced flow of reagent plus a thermal gradient and is referred to as the forced flow–thermal gradient (FCVI) process (Lackey and Caputo 1986, Lackey and Starr 1990). In order to force reagent to flow through the preform, a special holder is required as shown in Fig. 6 (Lackey

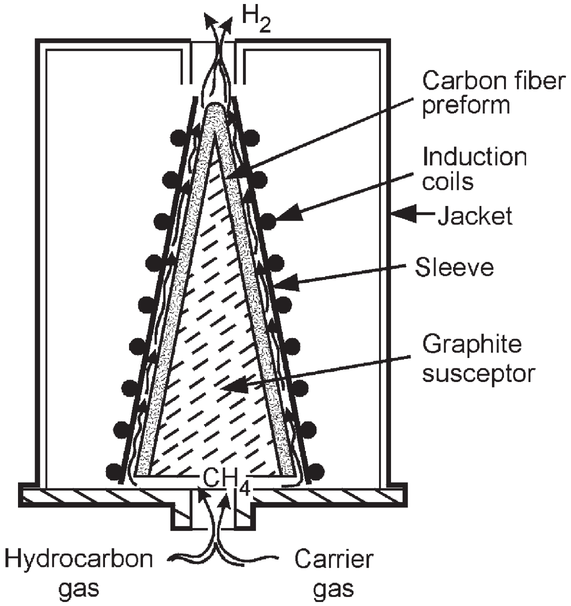


Figure 5
Conical preform being densified by the diffusion–thermal gradient CVI process (after Buckley and Edie 1993).

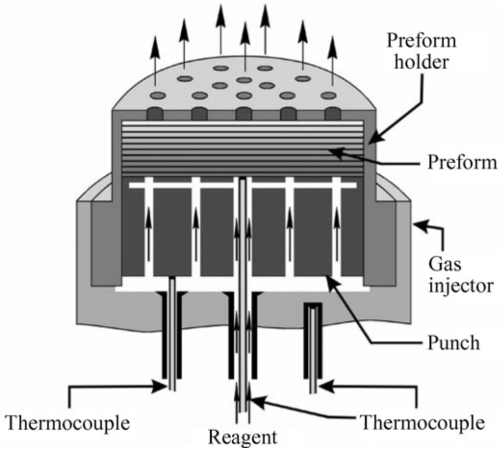


Figure 6
Apparatus for forced flow–thermal gradient CVI.

1996). The process is most highly developed for disks and tubes. In either event, the holder also permits cooling of one surface and heating of the other. The reagent stream first contacts the cooler surface and exits the hotter surface. Opposing factors permit densification to proceed uniformly throughout the thickness of the component, as follows. Slower deposition kinetics at the cooler surface favors less deposition than at the hot surface, but this is balanced by the fact that the higher reagent concentration at the cooler surface favors more deposition. In other words, some of the hydrocarbon is consumed as the reagent stream flows through the preform toward the hotter surface. By using experimentation and process modeling, it is possible to choose temperatures, temperature gradient, pressure, and reagent concentration and flow rate so that components with minimal spatial variation in density can be fabricated in less than five hours (Lackey and Vaidyaraman 1999, Vaidyaraman *et al.* 1996a, 1996b).

The three major processes for fabricating carbon–carbon have various advantages and disadvantages. The isothermal CVI process excels in that a large number of irregularly shaped parts can be placed in a single large furnace. Even through the processing time is very long, the cost per part is competitive. For the FCVI process, a disadvantage is the special fixture required for each part to direct the reagent stream to flow through the preform. The short processing time is an advantage and there is more freedom to select processing conditions (temperature, pressure, and reagent concentration) that yield the desired microstructure, and therefore properties of the deposited matrix. The inside-out process has intermediate capabilities compared to the isothermal and FCVI processes.

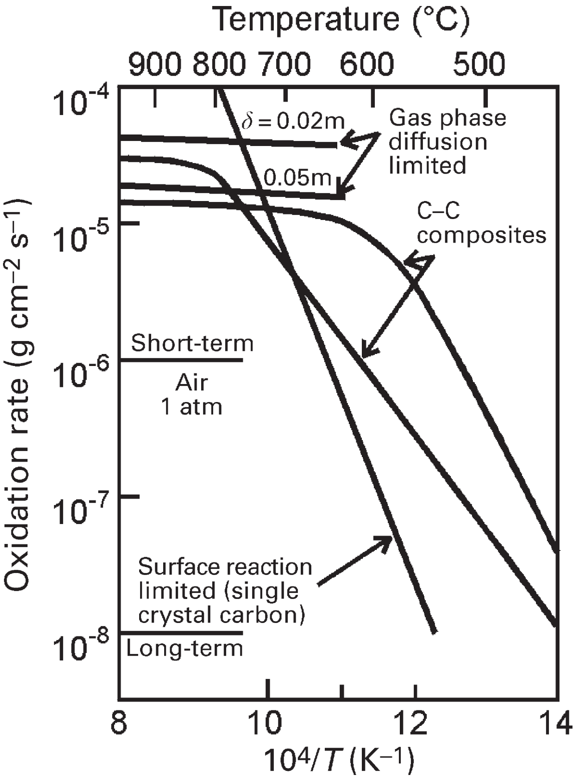


Figure 7
Calculated and observed oxidation rates for carbon–carbon (after Luthra 1988).

3. Oxidation Protection

Carbon begins to oxidize at a measurable rate in air at 370 °C, and the rate of oxidation increases rapidly with increasing temperature. In order for oxidation to occur, oxygen must diffuse inward through the gas boundary layer surrounding the component and the reaction product (CO or CO_2) must diffuse outward. Thus, either diffusion or reaction kinetics at the interface may be the rate-limiting step. Calculated and observed rates of oxidation for uncoated, uninhibited carbon–carbon composites have been compared (Luthra 1988). Figure 7 shows that for temperatures above about 700 °C the calculated values for gas-phase diffusion compare favorably with the experimental data. At lower temperatures, the experimental results are similar to the rates calculated for a surface reaction limited process. Since the oxidation rate at higher temperatures is controlled by gas-phase diffusion, the thickness of the boundary layer has a large influence on the oxidation rate. The boundary layer thickness (δ) is, of course, dictated by the test or in-service hydrodynamics. The rather thick boundary layers (0.02 m and 0.05 m) indicated in Fig. 7 are for samples contained in a crucible. Flow across a turbine blade would lead to a much thinner boundary layer. For a turbine blade, Luthra (1988) assumed a boundary layer thickness of 0.1 mm and calculated an oxidation rate of about $10^{-2}\,\mathrm{g\,cm^{-2}\,s^{-1}}$. Since acceptable rates are about $10^{-6}\,\mathrm{g\,cm^{-2}\,s^{-1}}$ and $10^{-8}\,\mathrm{g\,cm^{-2}\,s^{-1}}$ for short- and long-term exposures, it is clear that some means of retarding the oxidation of carbon–carbon is necessary for high-temperature applications in an oxidizing environment. This is an important aspect of carbon–carbon science and technology since failure of carbon–carbon components is often the result of oxidation of the matrix, fiber, or both. For this reason, oxidation protection approaches are described in some detail in this section.

3.1 Early Approaches

In 1934 the National Carbon Company patented a process for protecting carbon articles based on an inner layer of SiC and an outer glaze (glass) containing B_2O_3 (Johnson 1934). Glazes rich in P_2O_5 and SiO_2 were also mentioned in the patent. Many current patents are based on variations on this important theme. Throughout the 1940s and 1950s oxidation protection was sought for polycrystalline graphite to be used primarily as electrodes for melting metal or for furnace heating elements. In the 1960s oxidation- and erosion-resistant graphites were developed for reentry and rocket propulsion applications (Zeitsch 1967, Bortz 1971, Goldstein *et al.* 1966, Chown *et al.* 1963). These graphites contain boron, silicon, zirconium, and hafnium in the form of nonoxide ceramics, e.g., boron carbide. Oxidation produces a borate glass coating which is protective for periods of several hours at 1700 °C. Although primary protection by borate glasses can be expected for only limited times above ~1100 °C due to volatilization of B_2O_3, this early work demonstrated that incorporating boron into carbon articles was a valuable method for providing oxidation resistance.

The first effort to protect carbon–carbon composites from oxidation was conducted in the early 1970s for components to be used on the US Space Shuttle. These components were made of low-performance rayon-fiber-based composites (Rogers *et al.* 1975, 1976, Shuford 1984a). The coating system is composed of a SiC inner layer and an outer silicate glaze coating which contains SiC powder. A pack process in which the surface of the substrate reacts with silicon and SiO vapors is used to form the SiC layer. A silicon-rich coating is obtained, with the silicon particles being very small and finely dispersed. Coating thicknesses normally vary between 0.4 mm and 0.8 mm. Thicker coatings lead to decreased coating–substrate adherence and coatings over 2.5 mm thick are susceptible to spalling during heating. Glazing is performed with a commercial liquid silicate; cracks are sealed with silicon ethoxide (Shuford 1984a). Protection of low-strength, low-elastic-modulus rayon-based composites is easier than protecting high-performance composites because the thermal expansion of the rayon-based material is about twice that of high-performance carbon–carbons. Thus, the thermal expansion mismatch is not as severe. Modifications of this coating system containing boron have been developed for high-performance (nonrayon) carbon–carbon composites (Shuford 1984b).

Many types of coatings, other than glasses or glass formers, have been developed for protecting carbon-bearing articles from oxidation. As early as 1963 SiC coatings made both by CVD and reaction of the substrate surface with silicon were investigated, as well as a variety of carbide (e.g., ZrC) and boride (e.g., ZrB_2) coatings made by reaction sintering (Chown *et al.* 1963). These results showed the potential of CVD SiC for providing protection up to ~1700 °C; ZrC and ZrB_2 provided protection for short times to 2200 °C.

In the mid-1960s the use of iridium coatings to protect graphite was explored. Early work demonstrated excellent protection for short times at 2000–2100 °C. Iridium has a melting point of 2440 °C, is rather impermeable to oxygen to 2100 °C, is nonreactive with carbon below 2280 °C, and is an effective barrier to carbon diffusion (Strife and Sheehan 1988). Disadvantages of iridium are its excessive erosion due to volatile oxide formation, possible lack of adherence to carbon–carbon, and a relatively high thermal expansion coefficient. It is also expensive.

3.2 Protection Schemes

Basically three approaches for protecting carbonaceous materials from oxidation have evolved:

(i) coating with a solid material, e.g., SiC, iridium, etc.;

(ii) coating with glass; and

(iii) adding constituents which truly inhibit chemical reaction of carbon with oxygen or which react to form glasses which retard the ingress of oxidant.

In practice, all three approaches are often used simultaneously since each by itself is unsatisfactory except at low temperatures. Most of the so-called inhibitors act by forming a glassy layer over internal and external surfaces rather than truly inhibiting reaction between carbon and oxygen. Boron is thought to have a truly inhibiting effect on the reaction of carbon and oxygen; also, B_2O_3 is a glass former. Thus boron-based inhibitors can retard oxidation by both mechanisms. In addition, some inhibitors retard oxidation simply by acting as oxygen getters.

While the US Space Shuttle rayon-based carbon–carbon is protected by a glass layer applied over a SiC layer, the state-of-the-art approach for protecting high-strength carbon–carbon materials reverses the order of the coatings. That is, first a layer of glass (or

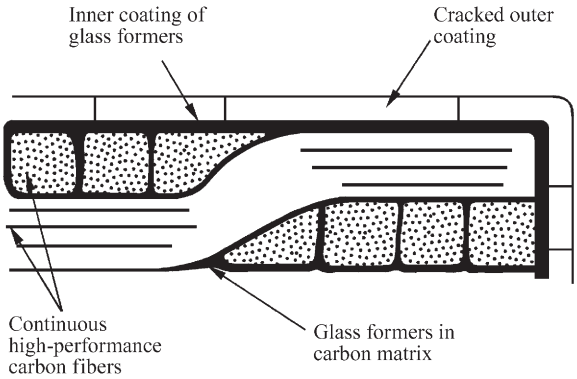

Inner coating of glass formers
Cracked outer coating
Continuous high-performance carbon fibers
Glass formers in carbon matrix

Figure 8
Oxidation protection scheme for advanced carbon–carbon composites.

glass formers) is applied and this is covered with an outer coating which is usually SiC but may be Si_3N_4 or some other material. This approach is shown in Fig. 8. The outer layer (e.g., SiC) provides the primary barrier to oxygen, but on cooling from high temperature, this layer cracks because of tensile stresses generated by thermal expansion mismatch with the substrate. The coating shrinks more than the major faces of a two-dimensional carbon–carbon composite. An additional source of stress, which in some instances is additive, is any as-deposited stress in the SiC layer.

The purpose of the glass layer located between the outer coating and the substrate is to provide a glass with sufficient fluidity to fill the cracks in the outer coating. The glass also coats the fibers and composite matrix in order to provide some in-depth oxidation protection. Limiting chemical reactions (glass formation) between the glass layer and the outer coating or oxide scales (e.g., SiO_2) formed on the outer coating is very important. The outer coating must not be aggressively attacked and the resultant glass must have suitable viscosity to permit effective sealing of cracks to continue. If the viscosity is too low, the glass may be blown from the composite by hydrodynamic forces present in most engine and aerospace applications. Formation of high vapor pressure compounds is undesirable since this leads to loss of the coating.

Ideally, the outer coating should retard the ingress of oxygen, i.e., it should be nonporous, oxidation resistant, and have a small diffusion coefficient for oxygen and other oxidants. Further, the coating should have a thermal expansion coefficient which is near that of the carbon–carbon composite. Unfortunately, most ceramics and metals have expansion coefficients which far exceed that for carbon–carbon (Strife and Sheehan 1988). Assuming that the outer coating is not an oxide, it will gradually oxidize. It is desirable that a protective oxide scale should form, i.e., the scale should have a low value for the diffusion of oxygen or gaseous reaction products. Scales of Al_2O_3, Cr_2O_3, and SiO_2 are noteworthy for their impermeability to oxygen. Although MgO is rather impermeable to oxygen, its large thermal expansion coefficient and the fact that its vapor pressure exceeds that of SiO_2 for temperatures below $\sim 1850\,^{\circ}C$ make it unattractive. Its susceptibility to hydration is also undesirable.

The major disadvantage of BeO is its toxicity. The combined relatively low thermal expansion coefficients and good oxidation resistance have led to the extensive use of silicon-based ceramics (SiC and Si_3N_4) for the outer coating. It is also significant that SiO_2 scales have low oxygen diffusion coefficients and can be modified with other oxides to control glass viscosity. Various oxides, nitrides, carbides, and borides have been considered as protective coatings. Coatings containing multiple phases and graded coatings have also been considered. Thermal expansion mismatch often proves to be troublesome. Coatings for providing protection to carbon–carbon for several thousand hours at cyclical temperatures in excess of $\sim 1300\,^{\circ}C$ are still under development.

4. *Microstructure of Carbon–Carbon*

The microstructures of both the fibers and matrix of carbon–carbon are important since they have a major influence on the mechanical, thermal, and electrical properties. Both the fibers and matrix can have a wide range of structures. The structures vary from amorphous to highly crystalline depending on the source of the material and the fabrication process, especially the maximum temperature used during processing. Important microstructural features include the degree of crystalline perfection, interlayer spacing, crystallite size, extent of preferred crystallographic orientation, and the quantity, shape, and orientation of pores. Extensive studies have created a large body of information on processing–microstructure–property interrelationships for carbon–carbon. Only a small fraction of that information can be summarized here. For further details see *Carbon Nanofibers**; *Carbon Mesophase*#; and Fitzer and Manocha (1998), Thomas (1993), Chung (1994), Savage (1993), and Buckley and Edie (1993).

4.1 Fiber Microstructure

The microstructure, and therefore properties, of carbon fibers varies markedly depending on whether the precursor is rayon, pitch, PAN, or CVD reagents, and on processing details. An important structural feature is the type and extent of preferred crystallographic orientation of the basal planes of the graphite-like crystallites. In many fiber types, there is a preference for the basal plane to be parallel to the long axis of the fiber with the *c*-axis oriented tangentially, radially, or randomly (Buckley and Edie 1993). The basal planes are often folded or crinkled, and "onion-skin" textures are common where the structure varies from the outside to the inside of a given fiber. Fibers with a large fraction of the basal planes parallel to the fiber length have high values for Young's modulus in the direction of the major axis of the fiber, since the bonds are strong within the basal planes.

The crystallite size parallel to the *c*-axis (L_c) and perpendicular to the *c*-axis (L_a) typically increases with increasing processing temperature. Typical values are in the range 2–20 nm. Fiber density typically increases with increasing processing temperature as a result of both a decrease in porosity and a decrease in the spacing between basal planes (d_{002}). The fiber structure influences not only the properties of the fiber but can also affect the properties of the carbon–carbon composite by influencing the adhesion between the fiber and matrix as well as influencing the crystallographic orientation of the matrix, especially near the fiber.

4.2 Matrix Microstructure

As with fibers, the matrix source (pitch, resin, CVI) and processing details control the microstructure of the matrix which, again like the fibers, can vary from amorphous to highly oriented graphite-like crystallites of varying density. Matrices obtained from thermosetting resins such as poly(furfuryl alcohol) and phenolics have minimum crystallinity which is not significantly enhanced by heating to typical graphitization temperatures.

Thermoplastic pitches attain their minimum viscosity at about 100–150 °C above their glass transition temperature. Low viscosity is advantageous for infiltration of the fiber preform. On further heating, volatilization of low molecular weight chemical species occurs along with polymerization and structural ordering of planar aromatic molecules. This liquid crystalline state is called a mesophase. Mesophase spheres about 0.1 μm in diameter form and subsequently coalesce on further heating. In the case of a pitch-derived matrix, the crystallinity in the final, solid carbon is spawned by the mesophase (Fitzer and Manocha 1998). The mesophase spheres deform and disk-like molecules of the liquid crystals tend to align parallel to the surfaces of fibers and fillers. This results in preferred crystallographic orientation near the fiber and filler. Within a bundle of fibers, this mechanism has a major influence on establishing the matrix microstructure. Further away from the fiber surfaces, the effect is reduced. The coalescence of the mesophase is retarded by the presence of particulate matter known as quinoline insolubles. These insolubles are found in coal tar pitches but not petroleum pitches. Therefore, petroleum pitches yield more highly crystalline carbon matrices. Carbonization pressure and temperature also influence the degree of crystallinity. Further, the rate of heating is important in that it influences whether evolved gases have time to escape or whether they form bubbles that eventually result in porosity in the carbonized product. Shrinkage, which accompanies carbonization and graphitization, causes the solid carbon matrix to crack. Formation of closed pores, which cannot be filled by subsequent liquid or gaseous infiltration steps, is undesirable if a high-density product is sought. Multiple infiltration–carbonization–graphitization steps clearly alter matrix microstructure as do oxidation–stabilization steps. Further complication arises when thermosetting resins and pitches are both used to produce the matrix. Additional details are available in Fitzer and Manocha (1998). Cracks and pores, which may cover the size range from macroscopic to microscopic, generally have a detrimental effect on mechanical properties.

A carbon matrix derived from a thermosetting polymer reflects, to some extent, the structure of the parent polymer. Such carbons are of low density (1.4–1.6 g cm^{-3}) and are not very graphitizable even at 3000 °C. They possess minimal crystal structure consisting of very small areas of short-range, parallel graphitic layers which are highly wrinkled. These carbons are typically isotropic because of the lack of preferred orientation. The presence of mechanical stress during carbonization can alter the degree of crystallinity. Variations in fiber–matrix bonding caused by different fiber surface characteristics can alter the stress state in the matrix near the fiber–matrix interface and thus alter the matrix crystallinity (Fitzer and Manocha 1998). As a result, the matrix near a fiber may consist of aligned graphitic layers while far away from a fiber the matrix may be isotropic. Just as in fibers, the graphite-like planes can be kinked. The presence of particulate fillers also influences the matrix microstructure.

The matrix of carbon–carbon prepared using any of the CVI processes is very dependent on the process details and is often described as being comprised of carbon having a rough laminar, smooth laminar, or isotropic microstructure. Rough laminar (RL) identifies carbon that is very anisotropic, i.e., high degree of preferred orientation of the basal planes, and readily graphitized. It is thought to occur when the growth rate is low (Thomas 1993). The material is dense and has well-defined growth features as shown in Fig. 1.

Smooth laminar (SL) carbon possesses less preferred orientation than RL and is less graphitizable. Isotropic (ISO) carbon, as the name implies, is even less oriented and most difficult to graphitize. As might be expected for a highly oriented, anisotropic material, RL carbon has the highest density, thermal conductivity, and elastic modulus (Rellick 1994).

Table 1
Apparent crystallite size, L_c, of heat-treated carbon and extinction angles for different microstructures of graphitized CVI carbon matrix.

Microstructure	Apparent crystallite size, L_c (Å)	Extinction angle (°)
RL	310–460	$\geqslant 18$
SL	95–165	4–18
ISO	70–110	<4

Techniques involving polarized light microscopy have been developed to rapidly identify the three types of material depending on the magnitude of the extinction angle as shown in Table 1 (Dupel *et al.* 1995).

Different sources are contradictory regarding identification of processing conditions that yield SL, RL, and ISO carbon deposits, and the present author does not believe it is possible to accurately generalize in any simple manner on this topic (Thomas 1993, pp. 91, 113, Fitzer and Manocha 1998, p. 151). For example, there is considerable disagreement as to whether low, intermediate, or high temperatures are conducive to the deposition of a particular microstructure. The disagreement is likely to be the result of complex interactions involving the processing variables as previously highlighted by Fisher (1993). That is, the influence on microstructure of one particular process variable, say temperature, is dependent on the levels of other process variables, such as pressure, flow rate, surface area, etc.

The processing variables of hydrocarbon type, concentration, flow rate, temperature, pressure, and preform architecture control in a complex manner the nature and residence time of numerous chemical species formed in the gas phase which, in turn, determines the microstructure of the deposited carbon matrix. Localized variations in structure can also occur as a result of the influence of the fiber surface. Also, deliberate or unavoidable changes in processing conditions during the course of infiltration/densification can alter the matrix structure. A thermodynamic microstructure model has been proposed that predicts the microstructure of the deposited carbon, which has been supported by experimental data in several different studies using isothermal and thermal gradient CVI (Lieberman 1972, Pierson *et al.* 1973, Pierson and Lieberman 1975, Lieberman and Noles 1973). The model has also been shown to be accurate for the FCVI process (Lewis *et al.* 1997). The model requires knowledge of the local temperature, pressure, and carbon-to-hydrogen ratio and is based on the assumption of equilibrium conditions in the gas phase before deposition of solid carbon. Regardless of the starting hydrocarbon, thermodynamic calculations show that under equilibrium conditions in the gas phase, methane (CH_4), hydrogen (H_2), benzene (C_6H_6), and acetylene (C_2H_2) are usually the most abundant gas species (Lieberman and Mark 1972). Gas chromatography results verified the calculations for thermal gradient CVI (Lieberman and Noles 1973). These results show that when methane and hydrogen are introduced into a furnace, the acetylene concentration initially increases, followed by an increase in benzene concentration. Carbon deposition is believed to result from the decomposition of both benzene and acetylene, although primarily from benzene. The Lieberman model does not attempt to identify the exact reaction mechanism for deposition. It simply uses the assumption that acetylene and benzene are important and that their relative concentrations determine the deposit microstructure. The value which is evaluated for the deposition conditions and used to predict the resulting microstructure is the gas-phase molar ratio, R, which is simply the equilibrium ratio C_2H_2/C_6H_6, i.e., the acetylene/benzene ratio. Experimental data show that as R increases, the microstructure changes from SL to RL to ISO. This is supported by the observation that deposited SL carbon is never seen directly adjacent to ISO carbon, and that a RL structure is always found between them. It has been observed experimentally that when R is below 5, the SL microstructure is obtained. When R is above 70, the ISO microstructure is obtained, and RL carbon is deposited for intermediate values (Lieberman 1972, Pierson *et al.* 1973).

Thermodynamic calculations show that the value of R has a strong dependence on temperature, pressure, and C/H ratios and perhaps accounts for the strong, complex influence these process variables have on the deposit microstructure. This model has been used to predict that it should be possible to use the FCVI process to fabricate carbon–carbon composites having the usually desired rough laminar matrix throughout the preform (Lewis *et al.* 1997). In this regard, the FCVI process offers considerable flexibility in controlling the matrix microstructure.

5. Properties

Fiber characteristics, preform architecture, the fiber–matrix interface, and the matrix macro- and microstructure control the properties of carbon–carbon. Addition of fillers, internal oxidation inhibitors, and an external oxidation protection coating also significantly influence mechanical, chemical, and physical properties of carbon–carbon. Obviously, a large number of material and process variables are involved, and caution must be exercised in specifying a particular carbon–carbon for a specific application.

5.1 Mechanical Properties

Most of the mechanical properties of carbon–carbon are very structure sensitive. This is particularly true of tensile strength and fracture toughness. As a result, it is not possible to simply list reliable values for these properties. Instead, the mechanical properties depend on whether rayon-, pitch-, or PAN-based fibers are used and whether resin, pitch, or CVI processes are used to form the matrix. Even if one specifies the fiber precursor type and consolidation process, the properties can still vary over a very wide range depending on the details of fiber and matrix processing, including postconsolidation thermal treatment and the nature of any oxidation protective coatings. Unfortunately, even if all of these details are specified, day-to-day upsets in the long series of fabrication steps can significantly alter the mechanical properties. This variability makes it necessary to test a large number of samples in order to obtain a reliable measure of the mean and distribution of mechanical property values. This lack of reproducibility, coupled with borderline fracture toughness and oxidation resistance, often makes designers reluctant to specify carbon–carbon for high-temperature, oxidative environment structural applications.

Strength, elastic modulus, and fracture toughness of carbon are improved by reinforcement with carbon fibers. Not surprisingly, the improvement is dependent on the volume fraction of fibers oriented in the test direction. As a result, the properties of, for example, a two-dimensional carbon–carbon are highly anisotropic with superior properties occurring in the plane of the lay-up, i.e., the x–y plane. Properties in the transverse direction (z-axis) are decidedly inferior.

For a unidirectional composite where all the fibers are oriented in a single direction, the maximum possible tensile strength can be calculated from the rule of mixtures assuming that the fibers and matrix are intimately bonded:

$$\sigma_c = \sigma_f V_f + \sigma_m V_m$$

where σ_c is the tensile strength of the composite in the direction of the long axis of the fibers, σ_f is the fiber tensile strength, V_f is the volume fraction of fibers, and σ_m is the matrix tensile strength (Chawla 1993). The fibers are usually so strong that the contribution of the matrix to the strength of the composite can be neglected, i.e., the composite strength is proportional to the fiber content. The strength calculated using the above equation exceeds that of actual composites by a factor of about two or more due to fiber damage, incomplete transfer of load to the fibers, stress concentrations, etc. (Savage 1993, pp. 290–2). Figure 9 shows for a unidirectional carbon–carbon that the strength decreases markedly when the angle between the applied load and the fiber axis is increased.

The strength of two- and three-dimensional composites can be estimated by taking into account the fiber loading in the direction of interest. As above, the actual strength is considerably lower than the theoretical maximum, an additional cause being misaligned fibers due to crimping caused by the weaving process.

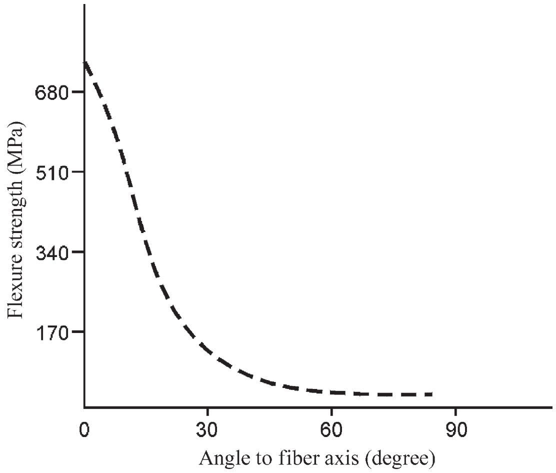


Figure 9
Effect of angle between fiber axis and loading direction on the strength of a unidirectional carbon–carbon (after Savage 1993).

Again, for a unidirectional composite, the Young's modulus of the composite can be estimated by the following equation:

$$E_{cl} = E_f V_f + E_m V_m$$

where E_{cl} is the Young's modulus of the composite in the longitudinal direction, E_f and E_m are the Young's modulus of the fiber and matrix, and V_f and V_m are the volume fraction of fibers and matrix, respectively (Chawla 1993). The above equations are derived assuming that the fiber and matrix are rigidly bonded, which is not at all accurate. In fact, it is well known that in carbon–carbon and other ceramic matrix composites, if the bond is strong, the composite will be very brittle. Therefore, considerable effort is directed at achieving a fiber–matrix bond of weak to intermediate strength. This contributes to fiber–matrix debonding, crack deflection, and fiber pullout which, along with matrix cracking, are important phenomena that significantly influence composite mechanical properties. As a result, the mechanics of fiber–reinforced composites is a complex topic and field of study in itself (Chawla 1993, Reddy 1997, Agarwal and Broutman 1990).

Figure 10 shows stress–strain curves for a series of two-dimensional carbon–carbon composites densified via CVI using 10% propane at temperatures between 1100 °C and 1400 °C (Oh and Lee 1988). The variations in strength, elastic modulus, and fracture energy with infiltration temperature could be caused by alterations of the strength of the fiber–matrix bond and variable matrix microstructure and composite density. Other process variables such as the hydrocarbon type and percentage of hydrocarbon, i.e., percentage of diluent, can have similar effects (Savage 1993, p. 298). Table 2 gives example mechanical property data for carbon–carbon made from pitch, PAN, and rayon fibers. Data for a grade of graphite are given for comparison.

Table 3 compares the mechanical properties of two resin-based carbon–carbon materials with that of a carbon–carbon having a CVI-derived matrix. Note that values are given for the as-carbonized as well as the graphitized materials.

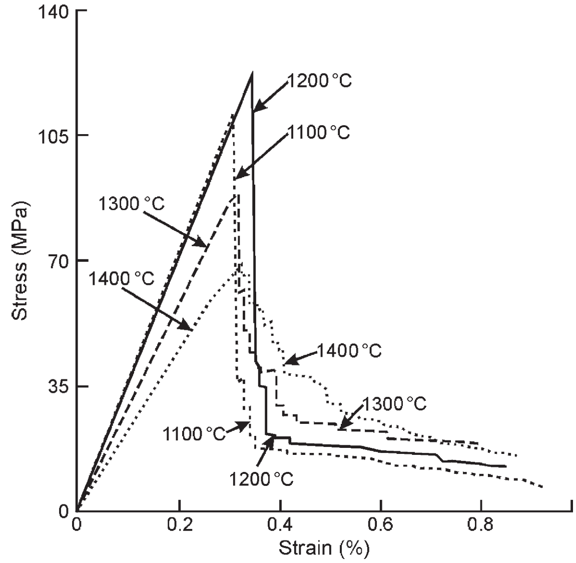


Figure 10
Room-temperature stress–strain curves for two-dimensional carbon–carbon prepared by CVI at 1100–1400 °C (after Oh and Lee 1988).

The strong influence of the fibers on the tensile strength of carbon–carbon is evident from examination of the data given in Table 4. Note that tensile strength is dependent on the quantity of fibers parallel to the loading direction. Also note the very low strength for the transverse direction. Compared to the other two, the three-dimensional composite is considerably more isotropic. It would also have increased strain tolerance, but maximum strength values are reduced.

It should be clear from the preceding discussion that the property values in Tables 2–4 are only representative and should be used with caution since processing details alter the values significantly. For a specific application, multiple pieces of material purchased from one or more vendors should be tested in an attempt to determine if the properties are suitable for the anticipated mechanical loading, including cyclic loading and environment. When materials such as graphite and carbon–carbon are selected for high-temperature applications, it is often because of the fact that these materials retain a significant fraction of their strength and stiffness at ultrahigh temperatures (~2500 °C) which are well above the maximum useful temperatures of other materials, including ceramics. The resistance to creep at high temperatures is excellent. Again, the exact influence of temperature on mechanical properties is dependent on the details of the manufacturing process. However, some carbon–carbons clearly show an increase in strength and stiffness as the temperature is raised to ~2000 °C (Thomas 1993). This may be the result of thermal expansion causing the closure of pre-existing cracks on heating or it may occur because the strength of some carbon fibers is known to increase with temperature. Other carbon–carbons show a minimum

Table 2
Mechanical properties of three carbon–carbon composites (A, B, C) and a graphite (G).

Property	Temperature (°C)	A[a]	B[b]	C[c]	G
Bulk density (g cm^{-3})		1.68	1.77	1.57	1.76
Young's modulus (GPa)		13.5	26.3	17.0	10.5
Flexural strength (MPa)		65.7	96.9		39.6
Tensile strength (MPa)	RT	35.7	55.4	68	28
	800	43.4	65.4	88	30
	1600	42.0	50.4	102	37
	2400	62.7	83.0	111	44
Fracture toughness (MPa m$^{1/2}$)	RT	2.96	3.44	4.0	0.8
	800	2.82	3.58	5.5	0.8
	1600	4.64	6.75	6.1	1.0
	2400	5.30	12.9	7.0	1.9
Thermal diffusivity (mm^2 s^{-1})		62.4	56.6		48.0

Source: Sato *et al.* (1989).
[a] Pitch carbon–carbon composite. [b] PAN carbon–carbon composite. [c] Two-dimensional rayon carbon–carbon composite.

Table 3
Density, porosity, and flexural strength of various types of carbon–carbon composites before and after graphitization.

	Resin-based matrix		
Property	With surface-treated fiber	With nonsurface-treated fiber	CVI matrix
Carbonized			
Bulk density ($g\,cm^{-3}$)	1.55		1.79
Open porosity (%)	4.5	7.0	13.0
Flexural strength (MPa)	150	900	520
Graphitized			
Open porosity (%)	4.5	10.5	12.0
Flexural strength (MPa)	600	350	500

Source: Kimura *et al.* (1989).

Table 4
Tensile strength (MPa) of one-, two-, and three-dimensional carbon–carbon.

Composite type	x-direction	y-direction	z-direction
Unidirectional fibers in x-direction	770	25	25
Bidirectional fibers in x–y plane	330	330	25
Three-directional, orthogonal weave; bundle size ratios x:y:z = 2:2:3	175	175	250

Source: Thomas (1993).

strength at ~1000 °C with higher strengths at lower and higher temperatures.

5.2 *Thermal Conductivity*

Some carbon–carbon composites have very high thermal conductivities. Since low density is important for some thermal management applications, a figure of merit defined as thermal conductivity divided by density may be used to rank materials. Table 5 shows such rankings, and it is apparent that carbon–carbon is much more suitable than traditional metals for many aerospace and electronic applications. This figure of merit is also pertinent for aircraft brakes. Carbon whiskers with thermal conductivities as high as $2500\,Wm^{-1}\,K^{-1}$ have been produced by vapor-growth techniques (Ting *et al.* 1995). Note in Table 5 that highly oriented CVD carbon is capable of a thermal conductivity very nearly as high as that of diamond. A carbon–carbon composite containing these fibers in a high-conductivity CVI carbon matrix thus offers the potential of an exceptionally high thermal conductivity, lightweight material. The thermal conductivity will be a maximum in the direction of the whiskers. Carbon–carbon having a thermal conductivity of $1000\,Wm^{-1}\,K^{-1}$ has been prepared (Lake *et al.* 1990). The microstructure of a high-conductivity, vapor-grown carbon whisker is shown in the high-resolution transmission electron micrograph of Fig. 2. Both thermal and electrical conduction in a material having the hexagonal graphite-like crystal structure is highly anisotropic since conduction within the basal plane is many times larger than that perpendicular to this plane. A highly crystalline carbon fiber or matrix possessing extensive preferred orientation of the basal planes will have high conductivity within this plane. Figure 2 shows that the basal planes in the vapor-grown fiber are oriented parallel to the long axis of the fiber and thus conductivity is very high along the fiber axis. The thermal conductivity of crystalline carbons, including carbon–carbon, increases with an increase in temperature from absolute zero, reaches a maximum conductivity between 0 °C and 500 °C, depending on the degree of crystallinity, and decreases with further increase in temperature (Savage 1993, p. 312). Higher graphitization temperatures clearly increase the thermal conductivity of carbon–carbon. The more easily graphitized RL carbon matrix structure yields higher conductivities than SL or ISO matrices. Pitch- and CVI-derived matrices yield higher conductivities than the less crystalline matrices derived from thermosetting resins.

5.3 *Electrical Conductivity*

The influence of processing on the electrical conductivity of carbon–carbon is similar to that previously described for thermal conductivity. That is, a

Table 5
Thermal conductivity and density of high-conductivity materials at 273 K.

Material	Thermal conductivity, K ($\mathrm{W\,m^{-1}\,K^{-1}}$)	Density, ρ ($\mathrm{g\,cm^{-3}}$)	Figure of merit, K/ρ
Carbon–carbon (in-plane)	~500[a]	2.0	250
Diamond (Type IIa)	2620	3.51	746
CVD carbon (in-plane)	2500	2.2	1136
Aluminum	236	2.70	87
Copper	403	8.91	45
Gold	319	19.31	17
Silver	429	10.5	41

[a] Depends on fabrication process.

carbon–carbon containing highly crystalline fibers and matrix has higher electrical conductivity and thus high graphitization temperatures, and the use of pitch or CVI matrices enhance electrical conductivity. After an initial increase, electrical conductivity decreases with increasing temperature. As a reference point only, the following values at room temperature are given for the electrical resistivity of a two-dimensional carbon–carbon. Parallel to the fibers, the electrical resistivity was about 35 μΩ m and 10 μΩ m for heat treatment temperatures of 1200 °C and 2800 °C, respectively, while perpendicular to the laminate the corresponding values were 106 μΩ m and 70 μΩ m (Huettner 1990).

5.4 Thermal Expansion

The thermal expansion of graphites and carbon–carbon is very anisotropic because of the vast differences in thermal expansion for the directions parallel and perpendicular to the basal planes of the hexagonal crystal structure. On heating from room temperature, there is an expansion in the *c*-axis direction, i.e., perpendicular to the basal planes, and actually a contraction in the *a*-axis directions, i.e., within the basal planes. Since both the carbon fibers and most carbon–carbon matrices exhibit preferred orientation of the basal planes, carbon–carbon exhibits extensive thermal expansion anisotropy. For the temperature range 25–1200 °C, a carbon fiber, Amoco T-300, has average thermal expansion coefficients of $8.85 \times 10^{-6}\,°\mathrm{C}^{-1}$ and $0.93 \times 10^{-6}\,°\mathrm{C}^{-1}$ for the radial and axial directions, respectively (Ellenburg *et al.* 1994). The axial value is about one-tenth that of a typical oxide ceramic fiber such as alumina. This low axial thermal expansion is a major reason for carbon–carbon having low thermal expansion coefficients within the plane of laminates, i.e., in the plane of the fibers of a one- or two-dimensional composite. In this direction, the thermal expansion coefficient of carbon–carbon containing high-modulus fibers is in the range $0–1 \times 10^{-6}\,°\mathrm{C}^{-1}$. In the transverse direction, the value is about $10 \times 10^{-6}\,°\mathrm{C}^{-1}$ (Thomas 1993, p. 208). The low expansion in the plane of the laminate severely complicates the application of an adherent external coating as discussed in Sect. 3.

6. Applications

Applications of carbon–carbon include military and commercial aircraft brakes, friction materials in clutches and synchronizers, leading edges of high-speed aircraft and re-entry vehicles such as the US Space Shuttle and intercontinental ballistic missiles, thermal management (heat sinks) in electronic devices, heat pipes, dimensionally stable space structures, molds for high-temperature forming of metals and glass, rocket nozzles, turbines, other heat engine components such as pistons and jet exhaust flaps, first walls in nuclear fusion reactors, corrosion-resistant containers, and various biomedical prostheses including hip and finger joints.

By far the largest application is for aircraft brakes. There are about 500 kg of carbon–carbon brakes in a typical commercial airliner. The low weight, high-temperature capability, high thermal conductivity, and desirable tribological properties of friction and wear rate are important. Carbon–carbon also has high internal friction, i.e., vibrations are readily dampened as a result of the presence of cracks and porosity. Vibration, which retards performance and can actually cause failure by fracture, is an important consideration for aircraft brakes.

The friction coefficient is another important property. It varies from more than 0.4 at taxi speeds to less than 0.1 for a rejected take-off. Moisture absorbed from the air and oxidation products formed during use can also influence the friction coefficient. Either pitch-or PAN-based fibers are used for brakes. A wide variety of preform types are used, including woven fabrics, short-length yarns, chopped fabrics, and knitted three-dimensional fabrics. Additional details regarding fabrication, design, and testing are available (Tatarzycki and Webb 1992).

Table 6
Costs for carbon–carbon.

Item	Percent of cost
Carbon fiber	17
Weaving	17
Matrix precursor	2
Fabrication	40
Protective coating	12
Pretax profits	12

Source: Savage (1993).

7. Cost

Cost is a major consideration for commercial aircraft brakes. The competing material is primarily steel. Carbon–carbon has superior heat capacity and thermal conductivity and it is lighter. However, carbon–carbon is expensive. Fitzer (1994) has estimated the cost of carbon–carbon as about US\$200 kg^{-1}. He further estimates that a little over 10% is associated with fiber and matrix raw material costs. The remaining costs are due to processing and presumably characterization. Savage (1993) states that because of the volume involved, aircraft brake materials are the cheapest carbon–carbon. He reports a price of £120 kg^{-1} for brake material, whereas a high-density, three-dimensional rocket nozzle may cost in excess of £1000 kg^{-1}. For a midprice-range (£300 kg^{-1}) carbon–carbon, he provides the breakdown of costs given in Table 6 (Savage 1993).

Bibliography

Agarwal B D, Broutman L J 1990 *Analysis and Performance of Fiber Composites*. Wiley, New York

Bortz S A 1971 Testing of graphite composites in air at high temperatures. In: Kriegel W W, Palmour H (eds.) *Ceramics in Severe Environments*. Plenum, New York, p. 49

Buckley J D, Edie D D 1993 *Carbon–Carbon Materials and Composites*. Noyes, Park Ridge, NJ

Chawla K K 1993 *Ceramic Matrix Composites*. Chapman and Hall, New York

Chown J, Deacon R F, Singer N, White A E S 1963 Refractory coatings on graphite, with some comments on the ultimate oxidation of coated graphite. In: Popper P (ed.) *Special Ceramics*. Academic Press, New York, p. 81

Chung D D L 1994 *Carbon Fiber Composites*. Butterworth-Heinemann, Newton, MA

Dupel P, Bourrat X, Pailler R 1995 Structure of pyrocarbon infiltrated by pure CVI. *Carbon* **33** (9), 1193–204

Ellenburg M G, Hanigofsky J A, Lackey W J 1994 Thermal stress analysis for coated fibers. *J. Mater. Res.* **9** (3), 789–96

Evans M J, Williams K A, Fisher R 1997 Manufacture of carbon fibre perform. *US Pat*. 5599603

Fisher R 1993 Manufacturing considerations for carbon–carbon. In: Thomas C R (ed.) *Essentials of Carbon–Carbon Composites*. Royal Society of Chemistry, Cambridge, UK, pp. 103–17

Fitzer E 1994 The future of carbon–carbon composites. In: Wright M A, Palmer K R (eds.) *Research Into Structural Carbons*. Southern Illinois University, Carbondale, IL, pp. 523–63

Fitzer E W, Geigl K H, Huttner W 1978 Studies on matrix precursor materials for carbon–carbon composites. In: *Proc. 5th Int. Carbon and Graphite Conf*. Society of Chemical Industry, London, Vol. 1, pp. 493–506

Fitzer E, Manocha L M 1998 *Carbon Reinforcements and Carbon/Carbon Composites*. Springer, Berlin

Girard H 1978 The preparation of high density carbon–carbon composites. In: *Proc. 5th Int. Carbon and Graphite Conf*. Society of Chemical Industry, London, Vol. 1, pp. 483–92

Goldstein E M, Carter E W, Kluz S 1966 The improvement of the oxidation resistance of graphite by composite technique. *Carbon* **4**, 273–9

Granoff B, Pierson H O, Schuster D M 1973 The effect of chemical vapor deposition conditions on properties of carbon/carbon composites. *Carbon* **11**, 177–87

Houdayer M, (Paris F R), Spitz J, Tran-Van D 1984 Process for the densification of a porous structure. *US Pat*. 4472454

Huettner W 1990 Potential of carbon/carbon composites as structural materials. In: Figueiredo J L, Bernardo C A, Baker R T K, Huttinger W (eds.) *Carbon Fibers, Filaments and Composites*. Kluwer Academic, Dordrecht, The Netherlands, pp. 275–300

Johnson H V 1934 Oxidation resisting carbon article. *US Pat*. 1 948382

Kimura S, Tanabe Y, Yasuada E. *Proc. 4th Japan–US Conf. Composite Materials*. Technomic, Lancaster, PA, pp. 867–74

Kotlensky W V 1971 A review of CVD carbon infiltration of porous substrates. In: *Materials '71, 16th Nat. SAMPE Symp. Exhibition*., Vol. 16, pp. 257–65

Kotlensky W V 1973 Deposition of pyrolytic carbon in porous solid. *Chem. Phys. Carbon* **9**, 173–262

Lackey W J 1996 Structural materials. In: Rees Jr W S (ed.) *CVD of Nonmetals*. VCH, New York, pp. 321–66

Lackey W J, Caputo A J 1986 Process for the preparation of fiber reinforced composites by chemical vapor deposition. *US Pat*. 4580524

Lackey W J, Starr T L 1990 Fabrication of fiber-reinforced ceramic composites by chemical vapor infiltration: processing, structure and properties. In: Mazdiyasni K S (ed.) *Fiber Reinforced Ceramic Composites*. Noyes, Park Ridge, NJ, pp. 397–450

Lackey W J, Vaidyaraman S 1999 Fabrication of carbon/carbon composites by forced flow-thermal gradient chemical vapor infiltration. *US Pat*. 5916633

Lake M L, Hickok J K, Brit K K, Begg L L 1990 Advanced Materials: Challenge Next Decade. *Proc. Int. SAMPE Symp. Exhibition*, Vol. 35. pp. 960–9

Lawton P G, Smith N 1992 Production of a shaped filamentary structure. *US Pat*. 5113568

Lewis J S, Lackey W J, Vaidyaraman S 1997 Model for prediction of matrix microstructure for carbon–carbon composites prepared by forced flow-thermal gradient CVI. *Carbon* **35** (1), 103–12

Lieberman M L 1972 Chemical vapor deposition of carbon: a model to relate gas phase conditions to structure of deposits. In: Glaski F A (ed.) *Proc. 3rd Int. Conf. on Chemical Vapor Deposition*. American Nuclear Society, Hinsdale, IL, pp. 95–119

Lieberman M L, Mark J L 1972 Report SC-DR-72 0775, Sandia Labs, Albuquerque, NM

Lieberman M L, Noles G T 1973 CONF 731017-3, Sandia Labs, Albuquerque, NM

Luthra K L 1988 Oxidation of carbon/carbon composites–a theoretical analysis. *Carbon* **26** (2), 217–24

McAllister L E Taverna A R 1972 Development and evaluation of Mod-3 carbon–carbon composites. In: *Proc. 17th Nat. SAMPE Symp.*, Vol. III-A-3

Oh S M, Lee J Y 1988 Effects of matrix structure on mechanical properties of carbon/carbon composites. *Carbon* **26** (6), 769–76

Olry P 1988 Process for manufacturing homogeneously needled three-dimensional structures of fibrous material. *US Pat.* 4790052

Pierson H O 1967 Development and properties of pyrolytic carbon felt composites. SC-DR-67-2969, Sandia National Laboratories

Pierson H O, Lieberman M L 1975 The chemical vapor deposition of carbon on carbon fibers. *Carbon* **13**, 159–66

Pierson H O, Lieberman M L, Smatana J F 1973 Report SLA-73-208, Sandia Labs, Albuquerque, NM

Pierson H O, Theis J D, Smatana J F 1968 Effects of deposition temperature on properties of carbon/carbon felt composites. SC-DR-68-264, Sandia National Laboratories

Reddy J N 1997 *Mechanics of Laminated Composite Plates: Theory and Analysis*. CRC Press, Boca Raton, FL

Rellick G 1994 Structure–property relationships in carbon/carbon composites. In: Wright M A, Palmer K R (eds.) *Research Into Structural Carbons*. Southern Illinois University, Carbondale, IL, pp. 313–38

Rogers D C, Scott R O, Shuford D M 1976 Material development aspects of an oxidation protection system for a reinforced carbon–carbon composite. In: *Proc. 8th Nat. SAMPE Tech. Conf.*, pp. 308–37

Rogers D C, Shuford D M, Mueller J I 1975 Formation mechanism of a silicon carbide coating for a reinforced carbon–carbon composite. In: *Proc. 7th Nat. SAMPE Tech. Conf.*, pp. 319–36

Rudolph J W, Purdy M J 1998 Pressure gradient CVI/CVD apparatus, process and product. *US Pat.* 5853485

Rudolph J W, Purdy M J 1999 Pressure gradient CVI/CVD apparatus, process and product. *US Pat.* 5900297

Rudolph J W, Purdy M J, Bok L D 1996 Apparatus for use with CVI/CVD processes. *US Pat.* 5480678

Sato S, Kurumada A, Iwaki H, Komatsu Y 1989 Tensile properties and fracture toughness of carbon-fiber felt reinforced carbon composites at high temperature. *Carbon* **27** (6), 791–801

Savage G 1993 *Carbon–Carbon Composites*. Chapman and Hall, New York

Sheehan P W, Liew R S 1996 Process for forming fibrous preform structures. *US Pat.* 5581857

Scaringella D T, Connors Jr. D E, Thurston G S 1996 Method for densifying and refurbishing brakes. *US Pat.* 5547717

Shuford D M 1984a Enhancement coating and process for carbonaceous substrates. *US Pat.* 4471023

Shuford D M 1984b Composition and method for forming a protective coating on carbon–carbon substrates. *US Pat.* 4465777

Smith N, Lawton P G 1995 Production of shaped filamentary structures. *US Pat.* 5388320

Strife J R, Sheehan J E 1988 Ceramic coatings for carbon–carbon composites. *Am. Ceram. Soc. Bull.* **67** (1), 369–74

Tatarzycki E M, Webb R T 1992 Friction and wear of aircraft brakes. In: *Friction, Lubrication and Wear Technology*. ASM International, Materials Park, OH, Vol. 18, pp. 582–7

Theis Jr. J D 1972 The process development and mechanical testing of a carbon/carbon composite fabricated by chemical vapor infiltration of a filament wound substrate. In: Glaski FA (ed.) *Proc. 3rd Int. Conf. on CVD*. American Nuclear Society, Hinsdale, IL, pp. 561–73

Thomas C R 1993 *Essentials of Carbon–Carbon Composites*. Royal Society of Chemistry, Cambridge, UK

Thurston G S, Suplinskas R J, Carroll T J, Connors Jr. D F, Scaringella D T, Krutenat R C 1998 Method for densification of porous billets. *US Pat.* 5733611

Ting J, Guth J R, Lake M L 1995 Light weight, highly thermally conductive composites for space radiators. *Ceram. Eng. Sci. Proc.* **16** (4), 279–88

Vaidyaraman S, Lackey W J, Agrawal P K 1996a Carbon/carbon processing by forced flow-thermal gradient chemical vapor infiltration (FCVI) using propane. *Carbon* **34** (5), 609–17

Vaidyaraman S, Lackey W J, Agrawal P K, Starr T L 1996b 1-D model for forced flow-thermal gradient chemical vapor infiltration process for carbon. *Carbon* **34** (9), 1123–33

Warren J W, Williams R M 1972 Isothermal CVD processing. In: *4th Nat. SAMPE Tech. Conf. Exhibition*, Palo Alto, CA, pp. 623–33

Zeitsch K J 1967 Oxidation-resistant graphite-base composites. In: Hove J E Riley W C (eds.) *Modern Ceramics*, Wiley, New York, p. 314

W. J. Lackey
Georgia Institute of Technology, Atlanta, Georgia USA

Ceramic Fibers from Polymer Precursors

High-strength, heat-resistant ceramic fibers can be produced from polymeric precursors by first spinning the precursor, and then transforming the polymer into a ceramic. This process is called the polymer precursor method.

The concept of the polymer precursor method was suggested by Chantrell and Popper (1964), and the synthesis of Si–C–N products from carbosilazane precursor was reported by Verbeek (1973). The polymer precursor ceramic fiber process was demonstrated by Yajima *et al.* (1975) with the development of a process for the production of SiC-based ceramic fibers by conversion of polycarbosilane. This type of fiber is produced by Nippon Carbon (NCK) under the trade name Nicalon. It is a continuous fiber with a diameter of about 15 μm. This fine fiber has a high tensile strength and has relatively high heat resistance. Another fine Si–Ti–C–O fiber from polytitanocarbosilane is produced by Ube Industries (UBE) under the trade name Tyranno. Besides Nicalon and Tyranno, a Si–C–N fiber from hydridopolysilazane (Legrow *et al.* 1987), a Si–N–O fiber from polycarbosilane (Okamura *et al.* 1987a), a high-purity Si_3N_4 fiber from perhydropolysilazane (Funayama *et al.* 1990), and a silicon carbonitride with high thermal stability (produced by Rhone-Poulenc under the trade name Fiberamic) have also been developed.

The SiC-based fiber Nicalon, produced by the thermal oxidation curing method (detailed below), has a high oxygen content and excess carbon, and is amorphous (Yajima *et al.* 1979). When Nicalon is heated above 1573 K, the fiber crystallizes into β-SiC, which is accompanied by the release of CO and SiO gases. The fiber shows severe degradation in tensile strength as a result of the structural destruction caused by the release of gases and the growth of SiC grains (Mah *et al.* 1984, Pysher *et al.* 1989, Laffon *et al.* 1989, Shimoo *et al.* 1990, 1994). A reduction in oxygen content is effective in improving the thermal stability of the SiC fiber. A curing method with electron beams has therefore been developed by Okamura *et al.* (1987b) in order to synthesize fibers with low oxygen content. When these fibers are converted from polycarbosilane to SiC under argon or nitrogen, the fiber obtained has excess carbon corresponding to a composition of about $SiC_{1.39}$; alternatively, if the conversion process is carried under hydrogen, the fiber has a near-stoichiometric composition of $SiC_{1.05}$ (Takeda *et al.* 1994). The resulting fibers are called Hi-Nicalon and Hi-Nicalon Type S, respectively. They have high heat resistance, the latter being excellent in this regard. By the heat treatment of Tyranno and its related Si–Al–C–O fiber, at higher temperatures in an inert atmosphere, sintered SiC fibers have also been produced. These are called Sylramic (Lipowitz *et al.* 1997) and SA (Ishikawa *et al.* 1998, Kumagawa *et al.* 1998) and exhibit excellent mechanical properties.

Other fiber systems can be produced using polymer precursors. For example, silicon carbide–silicon nitride ($Si_xN_yC_z$) fibers have been prepared by the pyrolysis of polycarbosilazane (Penn *et al.* 1983), silicon nitride fibers and various ceramic compositions containing the atomic species silicon, titanium, boron, nitrogen, and carbon have been prepared (Seyferth 1990), and Si–C–N filaments have been prepared from a new polycarbosilazane by melt spinning, γ-ray curing, and pyrolysis (Mocaer *et al.* 1993). However, the main engineering fibers derived by this route are based on silicon carbide. Nicalon and Tyranno are used as high-temperature fiber materials or as reinforcement fibers for polymer, metal, and ceramic matrix composites. More detail on the processing of these fibers is given in the following.

1. Synthesis of SiC Ceramic Fibers by the Polymer Precursor Method

Polycarbosilane (PCS) as a precursor polymer of SiC fibers is synthesized by the thermal conversion of polydimethylsilane. The molecular structure of PCS is planar, consisting of a ring-and-chain structure. PCS consists of two units, $Si(CH_3)_2CH_2$ and $Si(CH_3)HCH_2$, their unit ratio being almost the same (Hasegawa and Okamura 1986). The atomic ratio is $SiC_{1.93}H_{4.71}O_{0.01}$, oxygen being an impurity. The number average molecular weight is about 2023 and the melting point is 506 K. The SiC fiber processing sequence from PCS, shown in Fig. 1, comprises three main steps: PCS fiber spinning, fiber curing, and heat treatment of the cured fiber at high temperature.

The PCS is melt spun into a PCS fiber having a diameter of about 20 μm at about 583 K in N_2 gas. This PCS fiber is very fragile. It is therefore cured, in order to produce cross-links between polymer chains. This can be achieved by thermal oxidation in air at temperatures of 373–473 K or by electron beam (EB) irradiation, the dose rate varying from 2–5 $kGys^{-1}$ up to a dose of 15 MGy in helium gas. In thermal oxidation, as the PCS fiber consists of more than 7 mass% of oxygen, the PCS fiber becomes infusible by cross-linkage with Si–O–Si bonds. Under EB irradiation, Si–H and $-CH_2-$ or $-CH_3-$ groups of the PCS fiber are cross-linked and polymerized directly as shown in Fig. 2. After irradiation curing, the PCS fiber is heated to about 700 K for 1–2 h in order to diminish the free radicals trapped in the PCS fiber by irradiation. Then, the cured PCS fiber becomes stable in air and it is insoluble in solvents such as water and infusible for heating. The radiation-cured PCS fiber consists of less than 1 mass% of oxygen (Okamura and Seguchi 1992).

Alternatively, the PCS fiber can be irradiated with an EB in air at room temperature. The dose is 1–4 MGy with a dose rate $<2\,kGys^{-1}$. The oxygen content in the PCS fiber increases with the dose and can be controlled from a few to 30 mass%. The PCS

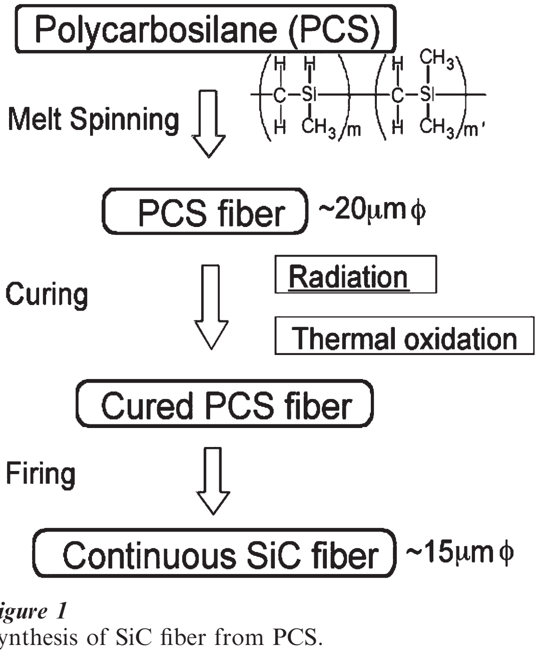


Figure 1
Synthesis of SiC fiber from PCS.

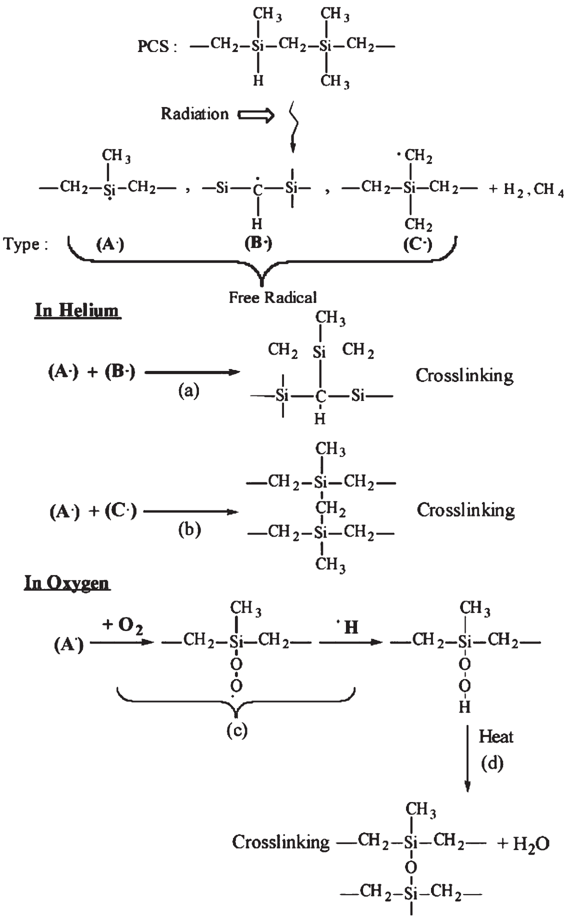


Figure 2
Mechanisms of radiation curing of PCS.

fiber is then cured by heat treatment at about 600 K for less than 1 h under vacuum after irradiation. This radiation/oxidation curing process is induced by cross-linking with oxygen bonding during the heat treatment.

The PCS fibers cured by these methods are finally heat treated in flowing argon or nitrogen at temperatures above 1473 K. This transforms the fragile as-cured PCS fiber to a high tensile strength SiC or near-SiC fiber. The SiC fiber obtained from the PCS fiber cured by thermal oxidation or radiation/oxidation (i.e., with oxygen cross-links) contains about 10 mass% oxygen, while the SiC fiber obtained from the PCS fiber cured with radiation cross-linking by EB irradiation contains about 1 mass% oxygen. The former fiber is termed Nicalon while the latter is called Hi-Nicalon. The Hi-Nicalon fiber has a carbon to silicon ratio of 1.39, but by heat treatment in flowing H_2 of radiation-cured PCS fiber, a SiC fiber having a stoichiometric composition can be prepared. This fiber is termed Hi-Nicalon-S.

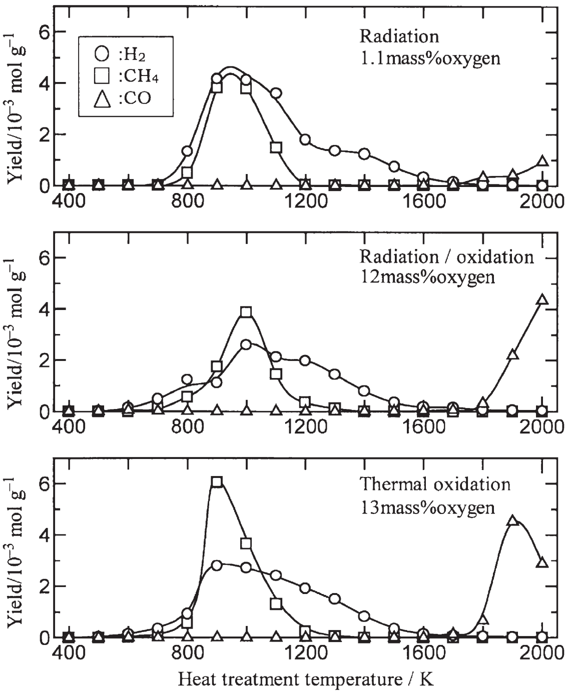


Figure 3
Gas evolution during heat treatment of cured PCS fibers.

2. *Conversion Process of the Cured PCS Fibers to SiC Fibers*

The conversion process of cured PCS fibers to SiC fibers is explained on the basis of the analysis of gases evolved during the heat treatment of cured PCS fibers (Sugimoto *et al.* 1995a). The gas evolution process of PCS fibers cured with EB irradiation, the radiation/oxidation process, and the thermal oxidation process are shown in Fig. 3.

In the case of the radiation-cured PCS fiber with 1.1 mass% of oxygen, H_2 and CH_4 are the major gases, which begin to evolve at around 700 K and increase to a maximum value at about 950 K. CH_4 evolution stops at about 1200 K, but the H_2 continues until 1700 K. The evolution of H_2 and CH_4 at temperatures of 800–1200 K corresponds to decomposition at the site of silicon atoms, i.e., Si–H and Si–CH_3; H_2 gas evolution at temperatures of 1000–1700 K is due to decomposition at the carbon atom sites in the PCS main chain (H–C–H). Above 1800 K, a small amount of CO gas is evolved.

In the case of the radiation/oxidation-cured PCS fiber with 12 mass% of oxygen, CH_4 evolution is almost the same as with 1.1 mass% of oxygen. H_2 evolution at 800–1200 K decreases to about 1/2, and the evolution at 1200–1700 K is nearly the same. Much CO gas evolves above 1800 K. This is due to the increased oxygen content in the cured PCS fiber.

In the case of gas evolution from the thermal oxidation-cured PCS fiber with 13 mass% of oxygen, CH_4 evolution is almost the same as that for the two processes mentioned above, but the peak is higher by about 50%. H_2 evolution is similar to that found after radiation/oxidation. A large amount of CO gas evolves above 1700 K; the increased CO gas evolution gives rise to a light green color and fine SiC particles grow on the fiber surface.

The PCS fiber cured by radiation/oxidation or thermal oxidation contains less hydrogen, because the hydrogen is lost by the formation of H_2O in the process of both oxidation-curing processes. The oxygen reacts preferentially with silicon atoms and results in a small amount of Si–H as shown in Fig. 2. Therefore, H_2 evolution at 800–1200 K decreases for the PCS fiber cured with oxygen in comparison with radiation curing. CO gas evolution also depends on the oxygen content in the PCS fiber, and is produced according to

$$SiC_xO_y \rightarrow SiC(s) + SiO(g) + CO(g) \qquad (1)$$

SiO and CO gas evolution accelerates the grain growth of SiC.

The conversion process of PCS to SiC also proceeds through free radical mechanisms (Sugimoto *et al.* 1995b). Some atomic sites are trapped in the pyrolyzed PCS fiber as free radicals, and detected by the measurement of electron spin resonance (ESR). The radicals are produced abruptly above 800 K and increase sharply in concentration with increasing temperature. In the case of the PCS fiber with low oxygen, the radical concentration reaches a maximum at 900 K, then decreases and shows a second peak at about 1300 K, and decreases to zero at 1700 K. For the PCS fiber with 12 mass% oxygen the maximum concentration is at about 1400 K, and the radicals remain until rather higher temperature than the low oxygen-cured PCS fiber. The trapped radicals reflect the state of the conversion process from the cured PCS fiber to SiC fiber. The presence of radicals indicates that the conversion process is progressing until 1700 K. Therefore, stable SiC fibers can be formed by heat treatment above 1700 K especially after EB irradiation.

3. *Properties of SiC Fibers*

The tensile strength increases during the conversion of the cured PCS fibers to SiC ceramic fibers as shown in Fig. 4. For the PCS fibers with 12 mass% oxygen and 13 oxygen, the synthesized SiC fibers reach maximum strength at 1600 K, and the strength decreases above this temperature. This decrease is due to the thermal decomposition described in Eqn. (1). By the generation of SiO and CO gases, the fiber morphology undergoes an extreme change, and the tensile strength rapidly decreases. For the PCS fiber with 1.1 mass% oxygen, the strength increases with the heat treatment temperature up to 1800 K. The SiC fiber formed from the PCS fiber with very low oxygen content of 0.4 mass% maintains its strength up to 2000 K. The Young's modulus (250 GPa) is maintained to 2100 K. By decreasing the oxygen content in PCS fibers with radiation curing, the fiber heat resistance is improved to 2000 K.

The thermal decomposition according to Eqn. (1) occurs in Tyranno fibers ($SiTi_{0.02}C_{1.37}O_{0.32}$) at temperatures above 1600 K. Large and loosely bonded SiC crystallites form, causing the fiber to become brittle. If, however, the fiber is heat treated with boron at high temperatures in an inert gas, the SiC grains in the fiber sinter together, thus eliminating porosity and resulting embrittlement. The chemical composition of this fiber is stoichiometric; this SiC fiber is called Sylramic. Another Tyranno-type fiber ($SiAl_{0.04}C_{1.47}O_{0.39}$) obtained from polyalunimocarbosilane is heat treated at temperatures above 1900 K in argon, and this, too, produces near-stoichiometric SiC sintered fibers. This SiC fiber is called Tyranno-SA. These sintered fibers have high tensile strength and high modulus. They show low creep rates, high thermal stability, and high oxidative resistance.

Several kinds of SiC fibers have been fabricated by the polymer precursor method, and their properties are shown in Table 1.

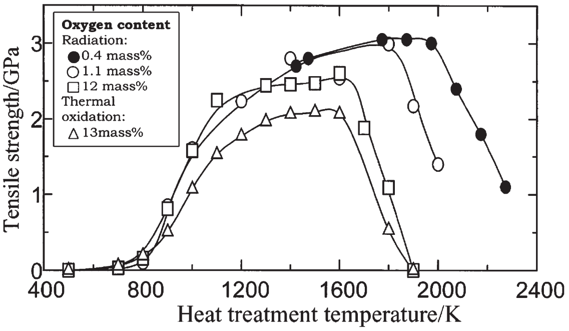


Figure 4
Tensile strength of SiC fibers produced by the heat treatment in argon from several kinds of cured PCS fibers. The tensile testing was carried out at room temperature.

Table 1
Properties of continuous SiC fibers.

	Nicalon			Tyranno		
	NL-200	HI-Nicalon	HI-Nicalon-S	Lox M	SA	Sylramic
Atomic composition	$SiC_{1.34}O_{0.36}$	$SiC_{1.39}O_{0.01}$	$SiC_{1.05}$	$SiTi_{0.02}C_{1.37}O_{0.32}$	$SiC_{1.08}$; O, Al < 0.008	$SiCTi_{0.02}\ B_{0.09}O_{0.02}$
Crystallite size (β-SiC, 111) (nm)	2.2	5.4	10.9	1.4	38	40–60
Tensile strength (GPa)	3.0	2.8	2.6	3.3	2.8	2.8
Tensile modulus (GPa)	220	270	420	187	420	400
Elongation (%)	1.4	1.0	0.6	1.8	0.7	0.7
Density (gcm^{-3})	2.55	2.74	3.10	2.48	3.02	>3.1
Diameter (μm)	14	14	12	11	10	10
Specific resistivity (Ω cm)	10^3–10^4	1.4	0.1	30		
Thermal expansion coefficient ($\times 10^{-6}K^{-1}$)	3.2 (298–773 K)	3.5 (298–773 K)				5.4 (293–1593 K)
Thermal conductivity ($Wm^{-1}K^{-1}$)	2.97 (298 K); 2.20 (773 K)	7.77 (298 K); 10.1 (773 K)	18.4 (298 K); 16.3 (773 K)		64.6	40–45

4. Conclusion

In the development process of ceramic fibers by the polymer precursor method, SiC-based fibers have mainly been produced from polycarbosilanes. The first fibers fabricated were amorphous; low-oxygen fibers were then developed; and finally sintered fibers were created. The high-temperature properties of these fibers have progressively been improved, and sintered fibers have excellent high-temperature resistance. The precursor polymers of these fibers are mainly polycarbosilane polymers prepared by the thermal decomposition of polydimethylsilane. In the near future, in order to obtain high-quality fibers with a diameter of 5 μm and to establish a low-cost fabrication process, more effective polymers as starting materials need to be developed.

Bibliography

Chantrell P G, Popper P 1964 In: Popper P (ed.) *Special Ceramics*. Academic Press, New York

Funayama O, Arai M, Tashiro Y, Aoki H, Suzuki T, Tamura K, Kaya H, Nishii H, Isoda T 1990 Tensile strength of silicon nitride fibers produced from perhydropolysilazane. *J. Ceram. Soc. Jpn.* **98**, 104–7

Hasegawa Y, Okamura K 1986 Synthesis of continuous silicon carbide fibre: 4. The structure of polycarbosilane as the precursor. *J. Mater. Sci.* **21**, 321–8

Ishikawa T, Kohtoku Y, Kumagawa K, Yamamura T, Nagasawa T 1998 High-strength alkali-resistant sintered SiC fibre stable to 2200. *Nature* **391**, 773–5

Kumagawa K, Yamaoka H, Shibuya M, Yamamura T 1998 Fabrication and mechanical properties of new improved Si–M–C–(O) fiber. *Ceram. Eng. Sci. Proc.* **19**, 65–72

Laffon C, Flank A M, Laridjani P, Hagege R, Olry P, Cotteret J, Dixmier J, Miquel J L, Hommel H, Legrand A P 1989 Study of Nicalon-based ceramic fibres and powders by EXAFS spectrometry, X-ray diffractometry and some additional methods. *J. Mater. Sci.* **24**, 1503–12

Legrow G E, Lim T F, Lipowitz J, Reaoch R S 1987 Ceramics from hydridopolysilazane. *Am. Ceram. Soc. Bull.* **66**, 363–7

Lipowitz J, Rabe J A, Zangvil A, Xu Y 1997 Structure and properties of SYLRAMIC silicon carbide fiber—a polycrystalline, stoichiometric β-SiC composition. *Ceram. Eng. Sci. Proc.* **18**, 147–56

Mah T, Hecht N L, McCullum D E, Hoenigman J R, Kim H K, Katz A P, Lipsitt H A 1984 Thermal stability of SiC fibres (Nicalon). *J. Mater. Sci.* **19**, 1191–201

Mocaer D, Pailler R, Naslain R, Richard C, Pillot J P, Dunogues J, Darnez C, Chambon M, Lahaye M 1993 Si–C–N ceramics with a high microstructural stability elaborated from the pyrolysis of new polycarbosilazane precursors: IV. Oxygen-free model monofilaments. *J. Mater. Sci.* **28**, 3049–58

Okamura K, Sato M, Hasegawa Y 1987a Silicon nitride fibers and silicon oxynitride fibers obtained by the nitridation of polycarbosilane. *Ceram. Int.* **13**, 55–61

Okamura K, Sato M, Seguchi T, Kawanishi S 1987b Application of electron beam irradiation for preparation of SiC fiber from polycarbosilane. In: Somiya S, Shimada M, Yoshimura M, Watanabe R (eds.) *Proc. 4th Int. Symp. Science and Technology of Sintering*, Tokyo, pp. 102–108

Okamura K, Seguchi T 1992 Application of radiation curing in the preparation of polycarbosilane-derived SiC fibers. *J. Inorg. Organomet. Polym.* **2**, 171–9

Penn B G, Ledbetter III F E, Clemons J M, Daniels T G 1982 Preparation of silicon carbide–silicon nitride fibers by the controlled pyrolysis of polycarbosilazane. *J. Appl. Polym. Sci.* **27**, 3751–61

Pysher D J, Goretta K C, Hodder R S Jr. Tressler R E 1989 Strengths of ceramic fibers at elevated temperatures. *J. Am. Ceram. Soc.* **72**, 284–8

Seyferth D 1990 Synthesis of some organosilicon polymers and their pyrolytic conversion to ceramics. In: Zeigler J M, Gordon Fearon F W (eds.) *Silicon-Based Polymer Science: A Comprehensive Resource*. Advances in Chemistry Series. American Chemical Society, Washington, DC, Vol. 224, pp. 565–91

Shimoo T, Chen H, Okamura K 1994 High-temperature stability of Nicalon under Ar or O_2 atmosphere. *J. Mater. Sci.* **29**, 456–63

Shimoo T, Sugimoto M, Okamura K 1990 Effect of crystallinity on pyrolytic rate of silicon carbide fibre. *J. Ceram. Soc. Jpn.* **98**, 1324–9

Sugimoto M, Shimoo T, Okamura K, Seguchi T 1995aa Reaction mechanisms of silicon carbide fiber synthesis by heat treatment of polycarbosilane fibers cured by radiation: I. Evolved gas analysis. *J. Am. Ceram. Soc.* **78**, 1013–7

Sugimoto M, Shimoo T, Okamura K, Seguchi T 1995bb Reaction mechanisms of silicon carbide fiber synthesis by heat treatment of polycarbosilane fibers cured by radiation: II. Free radical reaction. *J. Am. Ceram. Soc.* **78**, 1849–52

Takeda M, Sakamoto J, Imai Y, Ichikawa H, Ishikawa T 1994 Properties of stoichiometric silicon carbide fiber derived from polycarbosilane. *Ceram. Eng. Sci. Proc.* **15**, 133–41

Verbeek W 1973 Production of shaped articles of homogeneous mixtures of silicon carbide and nitride. Ger. Offen. 2218960

Yajima S, Hayashi J, Omori M 1975 Continuous silicon carbide fiber of high tensile strength. *Chem. Lett.* 931–4

Yajima S, Okamura K, Matsuzawa T, Hasegawa Y, Shishido T 1979 Anomalous characteristics of the microcrystalline state of SiC fibres. *Nature* **279**, 706–7

K. Okamura
Osaka Prefecture University, Osaka, Japan

Ceramic Matrix Composites with Roughly Equiaxed Reinforcements: Microstructure and Mechanical Behavior

The main goal of ceramic matrix composites is to obtain ceramic-based materials with improved mechanical properties, especially toughness and creep resistance. For this purpose, several reinforcement mechanisms can be used (Fantozzi and Olagnon 1993): phase transformation, microcracking, crack deflection, crack bridging, and fiber or whisker reinforcement.

These mechanisms will first be briefly presented. Then the application of these mechanisms on particle, whisker, and platelet-reinforced composites will be described.

1. General Background on Reinforcement Mechanisms in Ceramics

1.1 Transformation Toughening

Increases in strength and toughness can be obtained in ceramics that undergo a stress-induced transformation (Green *et al.* 1989). The case of ceramics containing zirconia particles that present a tetragonal to monoclinic phase transformation is the most spectacular. By using an addition of Y_2O_3, CeO_2, and others, a metastable tetragonal solid solution can be obtained at room temperature. Under the effect of an applied stress, a martensitic tetragonal (t) to monoclinic (m) transformation takes place in the vicinity of the crack, associated with a volume increase and a shear strain which impede the crack opening and therefore lead to an increase of toughness. This increase, ΔK_c^t, is given by (Evans 1984):

$$\Delta K_c^t = \frac{\eta E e^T V h^{1/2}}{(1-\nu)} \tag{1}$$

where η is a factor depending on the shape of the transformation zone around the crack, h its width, E the effective Young's modulus of the material, V the volume fraction of transformed particles, e^T the unconstrained transformation strain, and ν is Poisson's ratio.

The zone width and the volume fraction of transformable particles depend on the microstructure and especially on the grain size of the material, via the critical stress for transformation σ_c. The higher the grain size is, the lower is σ_c, thus the higher is the volume fraction of transformed particles; however, when the grain size becomes higher than a critical size, d_c, spontaneous transformation occurs upon cooling and most grains are already transformed before crack propagation.

1.2 Microcracking

When the size of the particles is greater than the critical size d_c, microcracking arises, which leads to a decrease in the strength of the material (Evans 1984). However if the particle size is smaller than the critical size, micro-cracks need additional stress to be induced that can be provided by the applied load. For an efficient reinforcement, the microcracking must be confined in a process zone in front of the crack tip. The particle size must be as close as possible to the critical size to induce significant toughening.

1.3 Crack Deflection

Cracks can be deviated either by the stress fields present around particles or by hard particles of higher resistance. This mechanism leads to a nonplanar crack, subjected to a mixed fracture mode. The analysis has been carried out for spherical, rod, and disc shaped particles for isotropic composite materials (Evans 1984). In the case of discs and rods, the influence of the aspect ratio has been investigated. High deflection and therefore high toughening are obtained with rods with a high aspect ratio for a volume fraction of about 10–20%.

This analysis has also been extended to the case of rods randomly oriented in planes in order to simulate transverse isotopic whisker-like composites. A more significant reinforcement can be obtained by crack deflection for this more specific case.

1.4 Crack Bridging

Such a reinforcement occurs when a crack meets a reinforcement particle without breaking it. The unbroken ligament exerts a bridging stress on the crack surfaces that counteracts the cracks opening and thus reduces the stress intensity factor (Becher and Rose 1994).

The toughening reinforcement is given by:

$$\Delta K_c^b = 2\left(\frac{V_f \sigma \mu u_c}{1-\nu}\right)^{1/2} \tag{2}$$

where V_f is the volume fraction of ligaments, σ the closing stress applied to the ligaments, μ the shear modulus, and u_c the crack opening at the end of the bridging zone.

The reinforcement depends on ligament density and the stress exerted by the ligaments, as well as on the crack opening. The reinforcement must exhibit a high strength in order to exert the highest possible stress and must be of large dimension in order to bridge over a large crack opening.

1.5 Whisker Reinforcement

The analysis of toughening by whisker reinforcement has been reviewed by Becher and Rose (1994). They consider uniaxially aligned whiskers. The analysis of reinforcement by whiskers is close to that developed for crack bridging. However, debonding of the whisker–matrix interface, and consequent friction at the interface, are generally taken into account. When whisker–matrix interface debonding occurs, the bridging stress supported by the whisker is reduced and a large bridging zone appears.

The energy dissipated by friction is given by:

$$\Delta G^f = \frac{V_f (\sigma_f^w)^3 r}{G E^w \tau_o} \tag{3}$$

where σ_f^w, E^w, r, and V_f are, respectively, the strength, the Young's modulus, the radius, and the volume fraction of whiskers, and τ_o is the interfacial shear stress.

The fracture energy increases as the interfacial shear stress decreases, i.e., when more interface debonding occurs. Furthermore, after whisker failure, whisker pullout can contribute to additional toughening.

In the case of short fibers, the same dependence of the fracture energy is observed: low fiber–matrix adhesion and high fiber strength are needed in order to obtain matrix microcracking and fiber pullout with friction and energy absorption.

2. Mechanical Properties of Zirconia-toughened Composites

As shown in Sect. 1.1, the fracture toughness increase is due to the t→m transformation of zirconia particles and needs the retention of t-ZrO_2. This retention can be obtained by using several types of microstructures (Claussen 1984). In all cases, the grain size of zirconia particles must be lower than the critical size d_c required for the spontaneous t→m. Zirconia can be found in the form of tetragonal zirconia polycrystal (TZP) stabilized in the tetragonal symmetry by an addition of different oxides, e.g., cerium (Ce–TZP) or yttrium (Y–TZP) oxides. Zirconia can also exist in the form of tetragonal precipitates in a cubic matrix. An example is that of magnesia partially stabilized zirconia (Mg–PSZ). Zirconia particles can also be dispersed in a ceramic matrix other than ZrO_2: these materials are called ZrO_2-dispersed ceramics. An example of such a composite is the ZrO_2-toughened alumina (ZTA) of which a microstructure is shown in Fig. 1.

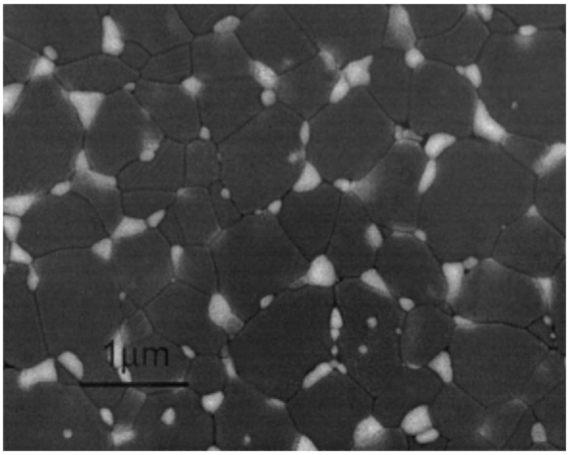


Figure 1
Microstructure of a zirconia (10 vol.%)-toughened alumina (ZTA). Zirconia appears as the bright phase in back-scattered scanning electron microscope observation.

2.1 Effect of Size and Amount of Zirconia Particles on Toughening

Toughness improvement is related to the tetragonal particle size. In the case of PSZ, the precipitate size depends on the ageing conditions after sintering and solution treatment. A dramatic drop-off in toughness occurs when the precipitate size is just above the critical size d_c: the t→m transformation becomes spontaneous during cooling and the amount of tetragonal precipitates available for transformation decreases. Figure 1 represents a ZTA composite where the zirconia grain size has been optimized in size (below d_c) and distribution (as narrow as possible) in order to obtain the largest amount of transformation toughening (Chevalier *et al.* 2000).

Equation (1) predicts that the toughness increment due to the t→m transformation depends on the volume fraction of transforming particles. This is actually observed (Becher and Rose 1994) when care is taken to control the tetragonal particles content. If during processing, monoclinic particles or stable (nontransforming) particles are obtained, the transformation toughening contribution decreases. Hence, fracture toughness presents a maximum with increasing volume fraction of zirconia particles. The maximum occurs at higher volume fraction with smaller particle size. This behavior is due to local tensile thermal stresses which provoke spontaneously transformation of the zirconia particles during cooling, decreasing the tetragonal phase content. Maximum toughness occurs for a volume fraction of particles corresponding to thermal stresses (which increase with the volume fraction of ZrO_2) reaching the critical stress σ_c. In this case, another contribution to toughening due to ZrO_2 particles that have spontaneously transformed is microcracking in the stress field around the monoclinic particles. Fracture energy (G_c) values for different pressureless sintered (PS) or hot-pressed (HP) ZTA compositions have been measured (Fantozzi and Orange 1986) (Table 1). For PS composites, G_c is nearly constant up to 15 vol.% ZrO_2 whereas, for HP materials, G_c shows a maximum near 10 vol.% ZrO_2.

The observed increase in G_c in PS materials and to 10% ZrO_2 in HP materials can be attributed mainly to the stress-induced transformation-toughening mechanism as indicated by the relative amount to tetragonal phase Qr and as predicted by Eqn. (1).

Above 15 vol.% pure ZrO_2, the relative amount of tetragonal phase is considerably reduced and microcracking toughening becomes the effective mechanism. However, the microcracking toughening is not as energy dissipative as transformation toughening, thus toughness is lower.

2.2 Effect of Solute Content and Temperature

With stabilizing agent additions, the initial tetragonal phase content increases (Table 1), which is reflected by an increase of the critical size d_c. Consequently, the stress-induced phase transformation toughening effect becomes preponderant over microcracking toughening and the fracture energy increases up to 2 mol.% Y_2O_3. For higher Y_2O_3 content, the critical stress σ_c required to initiate the transformation increases and the tetragonal phase becomes too stable to induce transformation toughening (G_c decreases).

Because the critical stress σ_c increases with yttria content and the thermal stresses augment with the volume fraction of ZrO_2, the volume fraction of particles corresponding to maximum toughness increases also as a function of yttria content.

Fracture toughness decreases as temperature increases, σ_c increases. The reinforcement becomes negligible when temperature reaches about 600 °C, the critical temperature at which the t→m transformation is no longer effective.

When the preponderant toughening mechanism is microcracking (e.g., the composite A20Z without stabilizing addition, in Table 1), toughening becomes independent of temperature.

Table 1
Fracture energy of ZTA ceramics.

Composition	A	A5Z	A10Z	A15Z	A20Z	A20Z1Y	A20Z2Y	A20Z3Y	A45Z3Y
ZrO_2 (vol.%)	0	5	10	15	20	20	20	20	45
Y_2O_3 (mol.%)	0	0	0	0	0	1	2	3	3
PS materials									
G_c (J m^{-2})	41	86	80	95	98	88	120	72	390
Qr (%)	0	98	48	20	15	33	95	98	96
HP materials									
G_c (J m^{-2})	50	160	250	115	102	117	270	194	576
Qr (%)	0	96	82	23	7	30	97	100	99

Source: Fantozzi *et al.* (2000).

Zirconia particles have also been added to reinforce other ceramic matrices. The preceding general features can be applied; however toughening mechanisms are more complex (Green *et al.* 1989).

2.3 Creep and Fatigue Behaviors of Zirconia-toughened Composites

The creep behavior of zirconia-reinforced ceramics has been described recently by Fantozzi *et al.* (2000). The addition of zirconia particles does not improve the creep resistance, creep being essentially controlled by grain boundary sliding (GBS) in the matrix.

The static fatigue behavior of ZTA composites was studied at room temperature (Chevalier *et al.* 2000). Figure 2 shows the slow crack growth behavior of an alumina reinforced under static loading. For comparison, the results obtained for alumina- and yttria-stabilized polycrystalline tetragonal zirconia (3Y–TZP) are also included.

Crack propagation is controlled by stress corrosion by water at the crack tip. Clearly, the slow crack growth resistance is increased for the ZTA composite, because the overall V–K_I (crack velocity versus stress intensity factor) curve is shifted towards high stress intensity factors for the composite. A threshold K_{IO} below which no propagation occurs is observed in the three ceramics. The threshold K_{IO} is higher for the alumina composite compared to alumina or TZP. This result is very significant, e.g., for bio-ceramics applications, where the resistance to slow crack growth is detrimental.

Cyclic loading promotes crack growth rate and leads to a decrease of the threshold stress. The fatigue behavior can be described as stress corrosion propagation associated with toughening degradation (as shown by Chevalier *et al.* 1999). Several toughening mechanisms can intervene (transformation toughening, bridging, microcracking). Similar observations have been made on PSZ composites by Dauskardt *et al.* (1990).

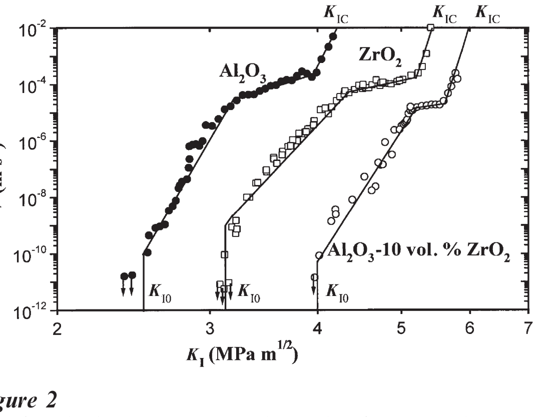


Figure 2
Crack velocity versus stress intensity factor K_I in alumina, zirconia (Y–TZP), and ZTA (10 vol.% ZrO_2) (after Chevalier *et al.* 2000).

3. Properties of Whisker Composites

Whiskers (W) being strongly anisotropic, different composite textures can be observed depending on the processing route used to produce the composite. The reduction of densification induced by the whiskers has led to the use of hot pressing (HP), which produces a microstructure in which the whiskers are randomly oriented within a plane. This is also the case when the slip-casting route is used. Isotropic bodies can be obtained by powder preparation followed with pressureless or hot isostatic pressing.

The structural anisotropy can be observed on the microstructure and by toughness and strength measurements. An increase of both toughness and strength of about 50% has been measured in SiC(W)–Al_2O_3 composites when the crack plane was perpendicular to the whisker plane (Becher and Wei 1984). Whiskers have been incorporated in alumina, silicon nitride and to a lesser extent in mullite, zirconia, glass, or other matrices (Bradley *et al.* 1988).

3.1 SiC(W)–Alumina Composites

Spectacular reinforcements by whiskers have been obtained in the case of alumina, with concomitant increases of both toughness and strength, Fig. 3. Toughness values higher than 10 Mpam$^{1/2}$ have been recorded for 20–30 vol.% whisker content. Such composites have been made with small diameter whiskers (0.3–0.6 μm). The achievement of a similar strengthening depends on the whisker dispersion in the matrix and on the densification of the composite. In order to reach theoretical density, the composites are generally hot pressed. Figure 3 shows the evolution of both

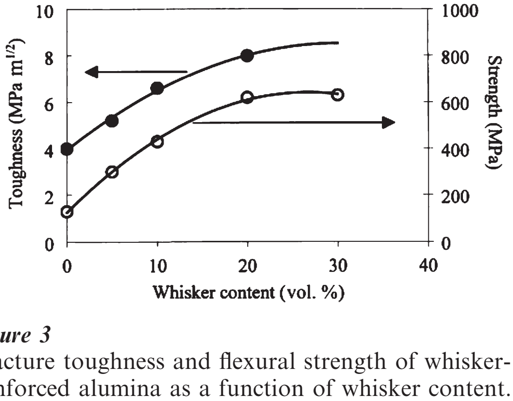


Figure 3
Fracture toughness and flexural strength of whisker-reinforced alumina as a function of whisker content.

toughness and flexural strength of alumina matrix composites with the SiC(W) addition.

As predicted by Eqn. (3), higher toughening is obtained with increasing whisker diameter, strength, and volume content, and with decreasing the frictional interfacial shear stress.

However, owing to the thermal expansion mismatch of SiC and Al_2O_3, too large a whisker diameter induces matrix microcracking, reducing strength. Optimum results have been obtained with whisker diameters lower than 1 μm (Fantozzi and Olagnon 1993).

The main toughening mechanisms are crack deflection and crack bridging. Reinforcement is observed up to about 1000 °C.

3.2 SiC(W)–Other Matrices Composites

Whiskers have been incorporated in silicium nitride, mullite, zirconia, and glass (Becher and Rose 1994). In the case of a Si_3N_4 matrix, the results were disappointing: the reinforcement is less significant than for alumina. This is due to the different microstructures observed for Si_3N_4: equiaxed or elongated grains depending on the sintering conditions and sintering aids. Toughness values of about 10 MPam$^{1/2}$ have been observed. For other matrices, the toughness increase is lower.

In order to improve toughness further, various toughening processes can be combined: whisker reinforcement, transformation toughening, and matrix microstructural-grain size effect. Either additive or coupled contributions can be obtained and multiple toughening mechanisms provide significant increase in fracture resistance: values of 10 MPam$^{1/2}$ have been observed for mullite reinforced by 20 vol.% of SiC(W) and 20 vol.% of t-ZrO_2 and toughness higher than 13 MPam$^{1/2}$ has been achieved for ZTA reinforced with 20 vol.% of SiC(W).

3.3 Creep Behavior

Most creep studies of these materials have been performed on whisker-reinforced alumina but other matrices such as Si_3N_4, ZrO_2, and mullite have also been studied (Fantozzi *et al.* 2000).

The creep rate of alumina is reduced by the additional SiC(W) as shown by Fig. 4. This decrease depends significantly on the whisker content, size and shape, on the presence of densification additives, and on the matrix grain size. The steady state creep rates of the alumina composites are two orders of magnitude lower than that of monolithic alumina. GBS, accommodated by diffusion processes, is the dominant mechanism responsible for the creep behavior of the composites. The addition of whiskers limits GBS, the whiskers acting as hard intergranular particles. For high whisker content ($\geqslant$30 vol.%), creep resistance is degraded, creep cavitation being enhanced. Similar enhancement in creep resistance is observed for mullite. However, reinforced Si_3N_4, ZTA, or ZrO_2 do not show improvement in creep resistance.

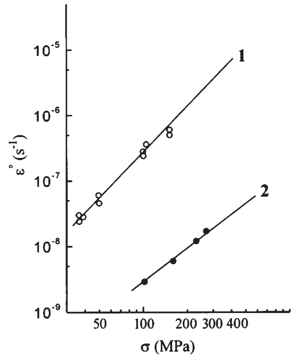


Figure 4
Stress dependence of the steady creep rate in flexure of alumina (1) and 20 vol.% SiC (W) composite (2) at 1473 K.

4. Particle and Platelet Reinforced Composites

Under this classification, only particles or platelets experiencing no transformation are considered.

The concept is similar to that of whisker composites, but with lower potential toughening. However, this type of composites has the advantage of being processed by conventional methods. The development of such composites was primarily focused on tribological applications with a double target: improving both toughness and hardness.

The reinforcement of alumina by SiC, TiC, BN, and TiN; of silicon carbide by TiB_2, TiC, and AlN; and of silicon nitride by SiC and TiC has been investigated (Fantozzi and Olagnon 1993).

The residual stress due to thermal mismatch is an important parameter controlling toughening. When the thermal expansion coefficient of particles is lower than that of the matrix, composites are likely to experience micro-crack toughening whereas in the contrary case, toughening will be mostly due to crack–particle interaction.

A good example of toughening by microcracking has been observed in SiC–Al_2O_3 composites: a maximum in toughness occurs as a function of the particle contents. Highest toughening is obtained with the smallest particles but a strength degradation is observed for large-sized particles, due to the formation of large cracks.

In the case of Si_3N_4 reinforced with SiC, the highest fracture energy has been obtained with the coarse particles, crack deflection being responsible for toughening without microcracking.

Thus, no general law between the toughness and the particle size can be stated since values observed strongly depend on the toughening mechanism. The reinforcement by particles can lead to values of 5–6 MPam$^{1/2}$ for toughness.

Platelets of SiC or Al_2O_3 have been incorporated into Si_3N_4, alumina, mullite, and TZP matrices. Toughening is due to crack deflection and crack bridging, with some pullout. In the case of Si_3N_4, spectacular toughening has been observed with 30% SiC platelets (14 MPam$^{1/2}$), but this was accompanied by a strength reduction. A special mention must be made with regard to nanocomposites (Niihara 1991). The reinforcing nano-sized particles are either inter- or intragranular. Indeed, room temperature and high temperature strengths are significantly improved in nanocomposites, although fracture toughness is only modestly increased. The particles bridging and the decrease of flaw sizes could be the essential strengthening mechanisms.

The creep behavior of particle or platelet reinforced composites has been studied for different matrices (alumina, mullite, Si_3N_4) (Fantozzi *et al.* 2000). In the case of alumina composites, the increase of creep resistance is lower than that observed with the whisker composites (one order of magnitude). On the other hand, alumina/silicon carbide nanocomposites present an excellent creep resistance: the creep rate is reduced by two to three orders of magnitude in comparison with alumina. The SiC nanoparticles hinder GBS and provoke a large improvement in creep resistance.

Improvement of creep resistance in other matrices for particle or platelet composites is also observed but is generally low. However, excellent creep resistance is obtained for mullite–zirconia composites and Si_3N_4 composites.

5. Conclusion

Monolithic ceramics can be toughened with roughly equiaxed reinforcements primarily to increase toughness, strength, creep resistance, and fatigue behavior. A wide variety of approaches can be used to reinforce ceramics: toughness values can be increased from 2–4 MPam$^{1/2}$ for monolithic ceramics to about 15 MPam$^{1/2}$ for composites with an optimized microstructure.

Bibliography

Becher P F, Rose L R F 1994 Toughening mechanisms in ceramic systems. In: Swain M V (ed.) *Materials Science and Technology: Structure and Properties of Ceramics*. VCH, Weinheim, Germany. Vol. 11, Chap. 8

Becher P F, Wei G C 1984 Toughening behavior in SiC-whisker-reinforced alumina. *J. Am. Ceram. Soc.* **67**, C-267–9

Bradley R A, Clark D E, Larsen D C, Stiegler J O 1988 *Whisker-and Fiber-Toughened Ceramics*. ASM International, Materials Park, OH

Chevalier J, De Aza A, Fantozzi G, Schehl M, Torrecillas R 2000 Extending the lifetime of ceramic orthopedic implants. *Adv. Mater*. (in press)

Chevalier J, Olagnon C, Fantozzi G 1999 Subcritical crack growth in 3Y–TZP ceramics: static and cyclic fatigue. *J. Am. Ceram. Soc.* **82**, 3129–38

Claussen N 1984 Microstructural design of zirconia-toughened ceramics. In: Claussen N, Rühle M, Heuer A H (eds.) *Advances in Ceramics: Science and Technology of Zirconia II*, American Ceramic Society, Columbus, OH, Vol. 12, pp. 325–51

Dauskardt R H, Marshall D B, Ritchie R O 1990 Cyclic fatigue-crack propagation in magnesia-partially-stabilized zirconia ceramics. *J. Am. Ceram. Soc.* **73**, 893–903

Evans A G 1984 Toughening mechanism in zirconia alloys. In: Claussen N, Rühle M, Heuer A H, (eds.) *Advances in Ceramics: Science and Technology of Zirconia II*, American Ceramic Society, Columbus, OH, Vol. 12, pp. 193–212

Fantozzi G, Chevalier J, Olagnon C, Chermant J L, 2000 Creep of ceramic matrix composites. In: Warren R (eds.) *Comprehensive Composite Materials, Vol. 4 Carbon/Carbon, Cement and Ceramic Matrix Composites*. Elsevier, Oxford, UK Chap. 7

Fantozzi G, Olagnon C 1993 Ceramic matrix composites. In: Chou T W (eds.) *Materials Science and Technology: Structure and Properties of Composites*. VCH, Weinheim, Germany, Vol. 13, Chap. 5

Fantozzi G, Orange G 1986 Thermomechanical properties of zirconia-toughened alumina materials. In: Moya J S, de Aza S (eds.) *Proc. Processing of Advanced Ceramics*, Sociedad Española de Ceramica y Violino, Arganda del Rey, Spain, pp. 187–215

Green D J, Hannink R H J, Swain M V 1989 *Transformation Toughening of Ceramics*. CRC Press, Boca Raton, FL, USA

Niihara K 1991 New design concept of structural ceramics—ceramic nano composites. *J. Ceram. Soc. Jpn.* **99**, 974–82

G. Fantozzi and J. Chevalier
INSA, Villeurbanne, France

Ceramic Matrix Composites: Applications

This article provides a short guide to the application of ceramic matrix composites (CMCs). Its purpose is to give the reader an appreciation for the broad range of potential for CMCs via illustrations of programs that have remained largely in the research and development phase, but which have demonstrated the directions for future applications.

The applications of CMCs are based on the fact that ceramics are key to meeting the "everyday" needs of society as well as enabling high-performance, futuristic concepts. Encompassing such a diverse range of materials as pottery, cements, high-temperature refractories and their coatings, and glasses, these materials are integral to our everyday lives.

Some of the well-known problems with ceramic materials, however, are that they are frequently fragile, lack toughness, and can fracture readily upon impact. Much of the technological history of ceramics has been devoted to overcoming these problems and notable successes are clearly evident. The highly impact-resistant window glasses found in automobiles and the use of steel rods to reinforce concrete are examples that are relied on every day.

1. General Application Requirements

As described in *Ceramic Matrix Composites: Matrices and Processing** and *Ceramic Matrix Composites with Roughly Equiaxed Reinforcements: Microstructure and Mechanical Behavior**, CMCs are a class of materials that address the issue of toughness by the incorporation of particles, fibers, and interfaces to inhibit the growth of cracks. The broad fields of building/construction, industrial equipment, and aerospace can all benefit significantly from the development of CMCs. The key attributes for successful application in these sectors can be summarized as follows:

Environmental durability. First and foremost, CMCs must exhibit the important environmental durability typical of ceramics. Polymers decompose and burn and metals corrode and melt under conditions that have little or no effect on many ceramics. This simple fact, combined with the relative low density of ceramics, ensures that CMCs will be well received in many potential applications, if they are durable and cost effective.

Strength combined with toughness. As described in *Continuous Parallel Fiber Composites: Fracture** and *Ceramic Matrix Composites with Roughly Equiaxed Reinforcements: Microstructure and Mechanical Behavior**, CMCs possess the unique capability of providing a combination of both strength and toughness (crack growth resistance) well in excess of traditional, monolithic ceramics. Because of the brittle nature of fracture in monolithic ceramics, toughness is the most important quality of CMCs. It is exemplified by the ability of CMCs to crack and deform locally, without causing total structural failure. This is necessary to provide reliability in the event that damage occurs during use or if loading in service exceeds the intended design limits.

Cost effectiveness. Finally, affordability must be achieved. In the development of CMCs this has been approached in two ways. First, CMCs are generally more costly than their unreinforced ceramic counterparts; however, CMC potential performance advantages and longer lives can "earn" this higher cost. Second, by virtue of the unique, new fabrication processes enabled by the structure of CMCs, it may be possible to manufacture very special new shapes or articles not even possible in ceramics. The following examples illustrate this concept.

2. Applications

2.1 Construction Industry Applications

The use of fibrous additives in brittle materials to facilitate their fabrication and use in construction applications is one of the oldest technologies known. The biblical reference to the use of straw in the manufacture of brick (Exodus 5:6) is a well-known example as is the more contemporary use of asbestos fibers in the manufacture of plasters and cements for domestic use. The application of cellulose fiber-reinforced plasters (papier-mâché) in Europe in the 1800s is an example of this technology being applied in decorative and practical ways. A comprehensive reference to the many ways in which reinforced cementatious materials can be created and used can be found in Hanant (1978).

Another example of achieving both toughening and strengthening is the use in Japan of small percentages (1–3 vol.%) of discontinuous carbon fibers in cement. The loosely bonded carbon fibers alter the fracture morphology to permit multiple microcracking and change the deformation behavior. This change in fracture characteristic causes the effective toughness of this material to become very high and difficult to quantify because cracks are diverted from their original path (i.e., no longer self-similar). Even at these very low volume percentages of fiber, the ability to alter the fracture mode can cause significant increases in reliability and structural strength. This technology has been so successful that it has led to the development of a lightweight concrete used in the construction of large buildings (Fig. 1).

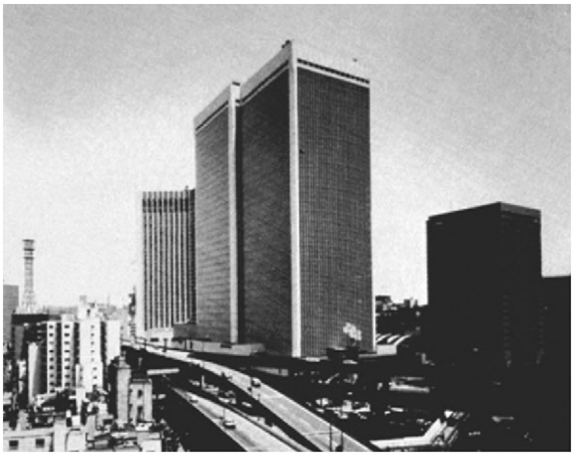


Figure 1
A 37 story building in Tokyo that has a carbon fiber-reinforced cement exterior. The low weight is achieved by adding large percentages of lightweight fly ash to the cement along with 2% chopped carbon fibers to increase toughness and strength.

2.2 Aerospace Applications

The impetus provided by potential aerospace applications has clearly been the key driver since the 1970s for the development of new types of CMCs. This has been a global effort with very large national teams leading the way. In Europe, Asia, and the USA there have been focused government-funded programs to apply ceramic composites to conventional gas turbine engine technology. The advantages sought have been lighter weight, improved durability, and greater efficiency as compared to metals-based designs.

(*a*) *Military fighter programs*. An example of a successful application is the European development of military aircraft components using a SiC matrix reinforced with either carbon or silicon carbide fibers. These composites are manufactured using the chemical vapor infiltration (CVI) process, as described in *Ceramic Matrix Composites: Matrices and Processing** and elsewhere (Naslain 1999). The application of these composites to the exhaust nozzles of high-performance military engines (Fig. 2) achieves mass savings of up to 60% compared with the more traditional metal components in service (Spriet and Habarou 1996).

(*b*) *Japan's Advanced Material Gas Generator Research and Development (AMG) project*. Japan, the country in which SiC-type fibers have reached their highest state of commercial development, also has major national programs to apply ceramic composites to gas turbines. Most notable is the AMG project, initiated in 1993, which involves the combined efforts of 14 major Japanese companies (Hiromatsu *et al.* 1995). The ambitious goals of this program include a 50% reduction in engine weight, a 20% improvement in fuel consumption efficiency, and a 75% decrease in the emission of nitrous oxides. The program has resulted in several key demonstrations of CMC performance. One example is a large CMC combustor liner, which operated without failure under conditions in which the gas temperature reached 1600 °C (Nishio 1998). Another example is a complex, rotating CMC gas turbine engine component known as a "blisk," a term to indicate that it consists of a series of blades (gas turbine airfoils) that are integrally embedded in a disk (Fig. 3). This significant transformation from metals-based gas turbine technology is attributable to the use of continuous fibers to make both the blades and the disk. The integration of these elements is accomplished via integral weaving of the entire structure, followed by composite densification to achieve the final shape (Nojima *et al.* 1999).

(*c*) *Supersonic commercial flight*. While the engine applications discussed above have clearly demonstrated the possibility for improvements in performance, it is the future concepts of supersonic commercial flight that capture the imagination and demonstrate the possibilities that CMCs can provide. A notable example of this is the national initiative in high-speed research in the USA. This effort is focused on the development of the technology to enable the creation of a 300 passenger civilian transportation aircraft that can travel 8000 km with a cruise speed of Mach 2.4. This would permit travel between Los Angeles and Tokyo in 4 hours, compared to the 10 hours of the late 1990s. An overview of this program

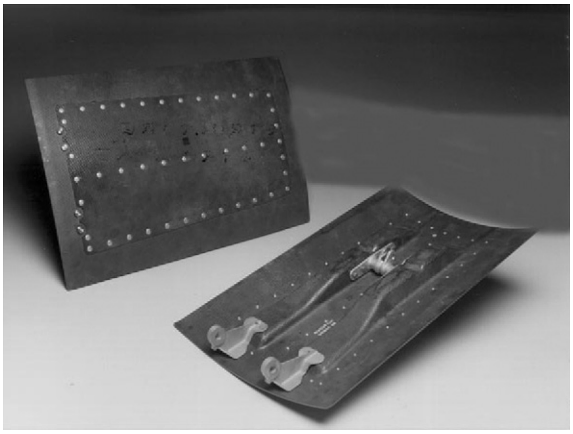


Figure 2
Carbon fiber-reinforced silicon carbide gas turbine engine exhaust nozzle flaps for the M88 engine of the Rafale fighter aircraft. The flaps are manufactured via the CVI process by SEP, Division of SNECMA, Saint-Medard en Jalles, France.

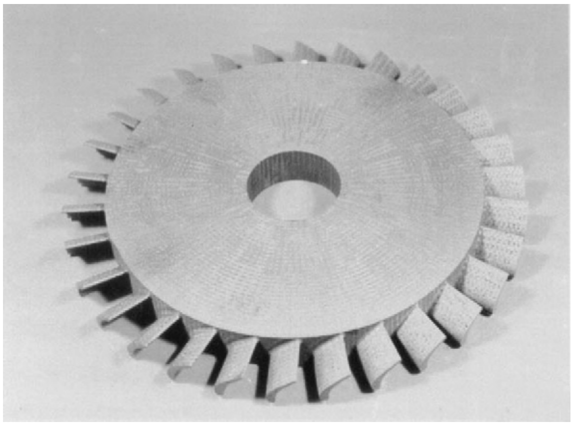


Figure 3
Integrally bladed gas turbine rotor disk (blisk) consisting of silicon carbide-type fibers in a silicon carbide matrix produced by CVI and polymer infiltration and pyrolysis. The rotor is approximately 250 mm in diameter and is manufactured as part of the AMG project, Japan.

can be found at the website www.lerc.nasa.gov/WWW/HSR/HSROver.html. The ability of CMCs to be formed into large, lightweight structural elements that can perform at high temperatures is key to the success of this aircraft and its ability to meet environmental requirements. Large CMC combustor liners (for reduced nitrous oxide emission) and nozzle acoustic liners (for reduced noise) have been successfully developed and tested under realistic engine conditions. Details of how CMCs have been used in these structures can be found at the website www.lerc.nasa.gov/www/hsr/epm.html.

This program has achieved significant milestones in the demonstration of durability and fabricability of complex structures. The impetus was curtailed when the US government decided to slow the pace of high-speed civil transport development; however, the materials technologies developed in this program will be used in other, more immediate applications in the gas turbine and industrial arenas.

2.3 *Industrial Applications*

While the aerospace industry has provided the technological basis for high-temperature CMCs, the cost of such CMC materials still remains high. Much larger volumes of production must be developed to bring costs down. Industrial applications for CMCs offer this opportunity.

The already widespread use of ceramics in many industrial applications provides a fertile basis for the future large-volume use of CMCs. A good source of information regarding existing ceramic applications is the American Ceramic Society (www.acers.org). The technical requirements to be successful in industrial applications (Frietag and Richerson 1998) are no less complex than those for aerospace applications. However, in an industrial environment prone to rough handling and occasional unanticipated damage, the fragile nature of ceramics frequently limits their applicability. This provides the opportunity for tough CMCs to enable application. As in the case of the construction industry, the need for achieving benefits while keeping the cost low is paramount.

A program dedicated to this expansion of CMC use into the industrial sector is the US Department of Energy-funded Continuous Fiber Reinforced Ceramic Composites (CFCC) program. Led by the US Oak Ridge National Laboratories, this program (www.ms.ornl.gov/programs/energyeff/cfcc/default/htm) addresses a broad range of applications, many of which are summarized in Table 1.

2.4 *Space-based Applications*

CMCs also find application where very specific properties can be developed via the creation of unique combinations of materials and microstructures. One such example is the development of dimensionally stable composites for space-based satellite applications. In this very special case, the development of a combination of dimensional stability, toughness, and thermal conductivity results in a material suited to the fabrication of both space structures and also mirrors. The latter are particularly demanding and are manufactured via the use of specialty glasses that require very tedious and expensive hot forming and machining operations to achieve their final, lightweight dimensions. In contrast, the incorporation of carbon fibers into a glass matrix can achieve similar levels of dimensional stability, with much higher thermal conductivity and stiffness and a significant reduction in weight. The mirrors shown in Fig. 4 are comprised of chopped carbon fibers in a borosilicate glass matrix. They have a thermal expansion coefficient of less than $0.1\,\mathrm{ppm}\,^{\circ}\mathrm{C}^{-1}$ from $-150\,^{\circ}\mathrm{C}$ to $+100\,^{\circ}\mathrm{C}$. Tests on board NASA's Long Duration Exposure Facility (LDEF) satellite have also demonstrated the extreme resistance of carbon fiber-reinforced glass composites to the effects of atomic oxygen (AO) degradation (Tredway and Prewo 1991).

Table 1
Industrial applications of CMCs under development.

Industrial application	CMC material type
Self lubricating diesel valve guides	Carbon fiber-reinforced silicon nitride produced via hot isostatic pressing
Tubular radiant burners and heat exchangers	Silicon carbide-type fiber-reinforced silicon carbide produced via chemical vapor infiltration
Hot gas filters and heat-treating furnace fan blades	Silicon carbide-type fiber-reinforced silicon carbide produced via chemical vapor infiltration
Immersion heating tubes for aluminum metal processing	Silicon carbide-type fiber-reinforced bonded silicon carbide
Land-based, power generation, gas turbine components (combustor liners, shrouds)	Silicon carbide-type fiber-reinforced silicon carbide produced via melt infiltration

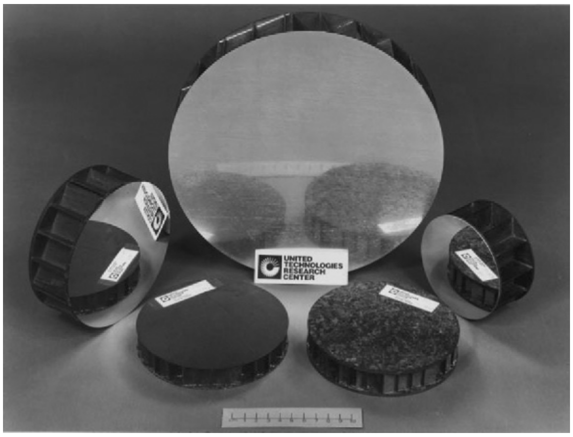


Figure 4
Mirrors for space-based optical applications. Manufactured with carbon fiber-reinforced glass, they exhibit extreme dimensional stability and highly polished surfaces.

3. Conclusion

The applications described in this article, along with many others achieved in a global community of national programs and companies, provide a robust basis for CMC technology readiness. The future realization of large-scale applications requires three things to happen. First, the need for CMCs must be shown conclusively; it must be demonstrated that metals and polymers alone will not suffice. Environmental requirements for cleaner and more efficient combustion will surely contribute to this clarification. Second, confidence in CMC production and performance must be established further. No serious applications will emerge without user confidence in the CMC suppliers and the performance of the resultant materials. Finally, CMCs must be affordable. The existent, well-recognized maturity and economic success of polymer matrix composites provides an excellent example that CMCs can follow.

Bibliography

Freitag D W, Richerson D W 1998 Opportunities for advanced ceramics to meet the needs of the industries of the future. Report NOE/ORO 2076, prepared for the US Department of Energy, available via the US Oak Ridge National Laboratory's website (www.ms.ornl.gov/programs/energyeff/cfcc/reports.htm)

Hanant D J 1978 *Fibre Cements and Fibre Concretes*. Wiley, New York

Hiromatsu *et al*. 1995 *Research and Development Status of Advanced Material Gas Generator (AMG) Project*, publication 95-*GT*-287 American Society of Mechanical Engineers, New York

Naslain R 1999 Materials design and processing of high temperature ceramic matrix composites: state of the art and future trends. *Adv. Composite Mater*. **8** (1), 3–16

Nishio K 1998 Development of the combustor liner composed of ceramic matrix composites. *Proc. Int. Gas Turbine and Aeroengine Congress*. American Society of Mechanical Engineers, New York

Nojima Y *et al*. 1999 Development of CMC bladed disk for advanced gas generator. *Proc. Int. Gas Turbine Congress*, Kobe, Japan

Spriet P, Habarou G 1996 *Applications of Continuous Fiber Reinforced Ceramic Composites in Military Turbojet Engines*, publication 96-GT-284. American Society of Mechanical Engineers, New York

Tredway W K, Prewo K M 1991 Fiber reinforced glass matrix composites for space structures. *Proc. 23 rd Int. SAMPE Technical Conf*., pp. 762–76

K. M. Prewo and W. K. Tredway
United Technologies Research Center, East Hartford,
Connecticut, USA

Ceramic Matrix Composites: Matrices and Processing

Ceramic matrix composites (CMCs) generally consist of ceramic fibers or whiskers in a ceramic matrix. CMCs are designed to overcome the main drawback of monolithic ceramics, namely their brittleness. They are referred to as inverse composites, which is to say that the failure strain of the matrix is lower than the failure strain of the fibers, whereas it is the reverse in most polymer or metal matrix composites. Hence, under load it is the matrix which fails first. In order to prevent an early failure of the brittle fibers when the matrix starts to microcrack, the fiber/matrix (FM) bonding should be controlled during processing. CMCs are tough materials and display a high failure stress when the FM bonding is not too strong or too weak, which is usually achieved through the use of a fiber coating referred to as the interphase (Chawla 1993, Warren 1992, Mazdiyasni 1990). The fabrication of CMCs requires specific processing techniques. Gas-or liquid-phase routes (or a combination of both), in which the interphase and the matrix are formed around the fibers from gaseous or liquid precursors, are usually preferred.

CMCs are used as thermostructural materials under severe service conditions, e.g., high temperatures under load and in corrosive atmospheres, such as combustion gases. The most commonly used CMCs are nonoxide CMCs, namely carbon/carbon (C/C), carbon/silicon carbide (C/SiC), and silicon carbide/silicon carbide (SiC/SiC), the fibers being specified first.

1. Structure

There are three constituents in a CMC: a fiber (or whisker reinforcement), a matrix, and generally an interphase. The nature, volume fraction, and arrangement of the constituents depend on a number of considerations including the function of the part to be fabricated, the service conditions, and the cost.

Both oxide and nonoxide fibers are used in practice. Refractory oxide fibers, such as alumina-based fibers, display a high resistance to oxidizing atmospheres and are insulating materials. However, their strength retention and creep resistance at high temperatures are poor due to grain growth. The most attractive oxide fibers are those made of a mixture of alumina and mullite ($3Al_2O_3$, $2SiO_2$), their service temperature limit being 1000–1100 °C. Nonoxide fibers include carbon and SiC-based fibers. With respect to oxide fibers, they exhibit a lower density and much better strength retention and creep resistance at high temperatures. The failure strength of carbon fibers increases with increasing temperature up to 1500–2000 °C and that of the best SiC fibers is still significant at 1400–1500 °C. Depending on their composition and microstructure, they are either conducting or insulating. Unfortunately all of them react with oxygen, above 500 °C for carbon fibers and above 800 °C for SiC fibers. The fiber arrangement depends on the load pattern. It is rarely unidirectional (1D) but rather multidirectional (nD, with $n \geqslant 2$). In this latter case, which is very common in CMCs, the starting material is a fiber preform, i.e., a self-standing porous body made of fibers (continuous or short) and having roughly the shape and size of the part to be fabricated (Bunsell and Berger 1999).

The interphase has a key role in a CMC although its volume fraction is extremely low. It is actually a thin coating (less than 1μm) deposited on the fibers or formed *in situ* during composite processing. It has three main functions: load transfer, crack deflection (the so-called mechanical fuse function), and diffusion barrier (in reactive FM systems). The best interphase materials might be those with a layered crystal structure (such as pyrocarbon (PyC) or hexagonal boron nitride (BN)) or a layered microstructure (such as $(PyC\text{–}SiC)_n$ multilayers, with $n = 5\text{–}20$) (Fig. 1(a)), the layers being oriented parallel to the fiber surface to promote crack deflection (Naslain 1998).

A variety of ceramic matrices are used. Silica-based glass–ceramics, i.e., ceramics made of a mixture of a crystalline refractory phase and a glass phase have been extensively studied (Mazdiyasni 1990). The most common are LAS, MAS, MLAS, CAS, and BMAS, with $L = Li_2O$, $A = Al_2O_3$, $S = SiO_2$, $M = MgO$, $C = CaO$, and $B = BaO$. All of them have two important advantages. First, before being ceramed, they are glasses which soften upon heating and display relatively low viscosities at high temperatures. This property is actually used to embed the fibers by hot pressing. Second, they exhibit a high resistance to corrosive media. However, the occurrence of residual glassy phase in the ceramed state lowers their creep resistance. Alumina and mullite, the latter having a better creep resistance than the former, are also used. However, their lack of plasticity at high temperatures is an important disadvantage in terms of CMC processing.

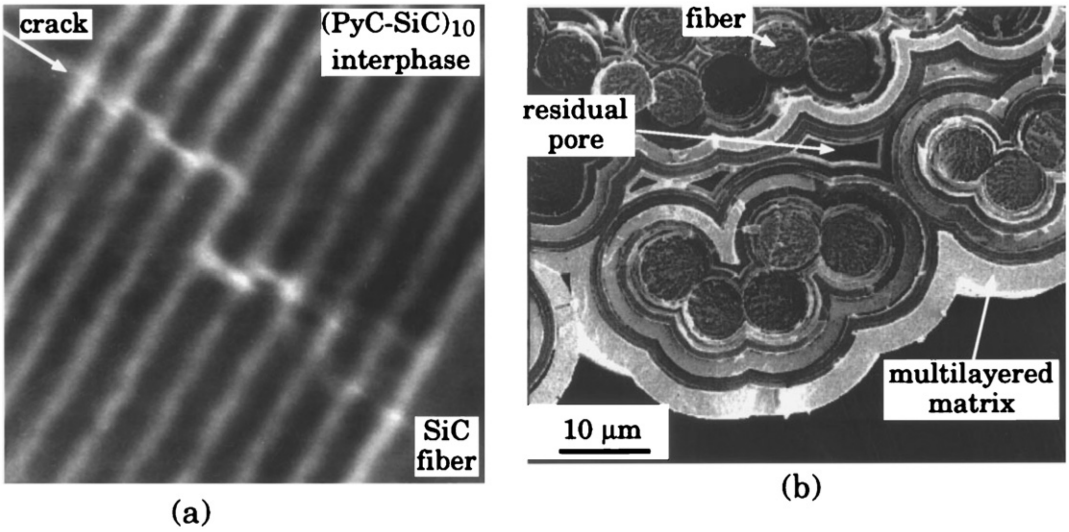


Figure 1
(a) Multilayered interphase and (b) self-healing matrix infiltrated in a fiber preform by P-CVI (courtesy of S. Bertrand and F. Lamouroux *et al.*).

Nonoxide matrices comprise mainly carbon, silicon carbide, and silicon nitride. Carbon matrices display a variety of properties depending on their microtexture. Pyrocarbon can be isotropic, nongraphitizing, and insulating on the one hand, or strongly anisotropic (laminar), graphitizable through heat treatment at very high temperatures (2500–3000 °C), and highly conducting (such as the so-called rough laminar used for C/C brake systems) on the other hand. Carbon is extremely refractory ($\approx 3600\,^{\circ}\mathrm{C}$), light ($1.8$–$2.1\,\mathrm{g\,cm^{-3}}$), and biocompatible (C/C composites are used in prosthetic devices) (Savage 1993). Silicon carbide is less refractory ($\approx 2500\,^{\circ}\mathrm{C}$) and its density slightly higher ($3.2\,\mathrm{g\,cm^{-3}}$). It is extremely hard and, more interestingly, its reaction with oxygen results in the formation of a protective layer of silica explaining why SiC ceramics can be used up to about 1500 °C in oxidizing atmospheres. Beyond this limit and/or at low oxygen partial pressure, the oxidation of SiC yields gaseous silicon monoxide (SiO) with no longer any protective effect. SiC matrices can actually be pure SiC or mixtures of SiC with free silicon or free carbon. When free silicon is present, the service temperature limit is lowered to $\sim 1410\,^{\circ}\mathrm{C}$. The occurrence of free carbon does not alter the refractoriness of SiC matrices. Silicon nitride matrices have properties close to those of SiC matrices; however, they are less thermally stable and exhibit lower conductivities.

When selecting a given fiber–interphase–matrix system, several considerations should be taken into account. First, the constituents should be compatible. The coefficients of thermal expansion (CTEs) of the fiber, α_f, and the matrix, α_m, should be roughly the same. If $\alpha_m > \alpha_f^R$ (the radial CTE of the fibers), the matrix clamps the fibers after cooling from processing to ambient temperature, resulting in strong FM bonding and a brittle character. Conversely, if $\alpha_m < \alpha_f^R$, the FM interface may spontaneously debond during cooling. The matrix should also be chemically compatible with the fibers. FM reactions may alter the strength of the pristine fibers and/or change the FM bonding. As an example, SiC fibers react with silica-based glass–ceramic matrices and a suitable interphase, such as a BN–SiC dual interphase, has to be used. To avoid these problems, the fibers and the matrix often have relatively close compositions in actual CMCs.

CMCs should also be compatible with their service atmospheres. Mostly, these service atmospheres are oxidizing (air or combustion gases) whereas the best CMC constituents are nonoxides, a somewhat paradoxical situation. To solve this key problem, two routes have been suggested: (a) the use of oxidation inhibitors and protective coatings and (b) the design of self-healing interphases and matrices. The addition of boron or phosphorous compounds (inhibitors) to carbon matrices decreases significantly their oxidation rate. Multilayered coatings containing glass formers (such as silicon or boron carbides) are also used to slow down the in-depth diffusion of oxygen in C/C composites. The concept of multilayered self-healing interphases (Fig. 1(a)) and matrices (Fig. 1(b)) has been developed for advanced SiC matrix composites, with a dramatic increase in their lifetime in oxidizing atmosphere at high temperatures and under cyclic loading (Lamouroux *et al.* 1999). Here, the interphase and/or the matrix, deposited according to a gas-phase route, are multilayered, involving crack arrester layers (such as PyC or BN) and glass former layers (such as B–C, Si–C, or Si–B–C compounds). When a microcrack is formed, oxygen diffusing along that crack reacts with the boron- or silicon-bearing species, yielding a low-viscosity glass which fills the crack and slows down oxygen diffusion. Under such conditions, the composite behaves somewhat as a smart material. It displays high mechanical performance with a lifetime long enough for applications under oxidizing atmospheres (e.g., in engines) (Naslain 1999).

2. *Processing Routes*

Most ceramic matrices being hard, refractory, and displaying no (or little) plasticity within a broad temperature range, the brittle fibers are preferentially embedded in a liquid or gaseous precursor of the matrix which is transformed into a ceramic matrix via chemical reactions or annealing treatments (Mazdiyasni 1990, Naslain 1999, Warren 1992).

2.1 Liquid-phase Routes

(*a*) *Reactive melt infiltration* (*RMI*). RMI processes can be used when one of the elements constituting the ceramic matrix displays a low (or relatively low) melting point and readily wets the fibers. Such is the case for aluminum alloys (m.p. $\approx 650\,^{\circ}\mathrm{C}$) and even silicon (m.p. $= 1410\,^{\circ}\mathrm{C}$). The former can yield an alumina matrix by reaction with an oxidizing atmosphere and the latter a silicon carbide matrix through a chemical reaction with a carbonaceous fiber preform.

In the so-called DIMOX process (from a company called Lanxide, DE, USA), the molten element, aluminum, reacts with the atmosphere, e.g., air, as it climbs by capillary forces in the preform pore network, yielding an alumina matrix. The alumina matrix grows according to a complex mechanism involving the dissolution of oxygen and the precipitation of alumina, in which magnesium (present in the alloy) plays a key role. A layer of magnesia is formed at the metal surface which actually prevents the formation of a dense protective scale of alumina that would otherwise stop the in-depth diffusion of oxygen. The process being run at about 1200 °C, a variety of fibers, including Si–C–O Nicalon fibers (from Nippon Carbon), can be used. The process can be extended to nitride matrices (such as aluminum nitride and even silicon nitride) if it is conducted in a nitrogen atmosphere (Mazdiyasni 1990, Urquhart 1991).

In the so-called SILCOMP (from General Electric) and related processes, a ceramic fiber preform is first consolidated with carbon, e.g., according to the PIP process (see below). Then it is impregnated with liquid silicon at about 1450 °C (Fig. 2). Liquid silicon spontaneously wets carbon and hence climbs by capillary forces in the preform pore network, reacting with the pore wall, i.e., carbon, to give the SiC matrix (Mazdiyasni 1990, Luthra *et al.* 1993). Since the process is run at high temperatures, only fibers with a high thermal stability, namely carbon or oxygen-free SiC fibers, can be used.

Both processes are relatively rapid and yield near-net-shape composites. Conversely, liquid aluminum and silicon being corrosive, the fibers have to be precoated. Dual BN–SiC coatings deposited from gaseous precursors (see Sect. 2.3) have been successfully used in Nicalon/alumina and Hi-Nicalon/SiC composites. The BN sublayer acts as mechanical fuse (crack deflection function) and the SiC sublayer as diffusion barrier. Further, some residual free aluminum or silicon is usually present in the matrix as the result of an incomplete chemical reaction, which lowers the mechanical properties at high temperatures. Residual aluminum can actually be removed by wicking (thus yielding some open porosity) whereas residual silicon can be entrapped in the SiC matrix as a refractory silicide when necessary.

(*b*) *Polymer impregnation and pyrolysis* (*PIP*). In the PIP process, a porous fiber preform is first impregnated with a liquid organic (or organometallic) precursor of the matrix. After curing by heating, irradiation, or chemical reaction, the material is pyrolyzed at medium temperatures (700–1000 °C). When necessary, the composite is further submitted to a heat treatment (2500–3000 °C for carbon) to stabilize its microstructure and optimize its properties (Fitzer and Manocha 1998, Mazdiyasni 1990, Savage 1993). The liquid precursor is used either in the molten state or in solution. Examples of commonly used precursors are thermosetting polymers (e.g., polyphenolics) or mesophase pitch for carbon, polycarbosilanes for SiC, and polycarbosilazanes for Si–C–N matrices. The liquid phase should display a low enough viscosity and wet the fibers (in the impregnation step).

Further, the precursor should have a high ceramic yield, typically 70–85 wt.% (in the pyrolysis step). Even under such conditions a weight loss and shrinkage of the matrix occur during the pyrolysis due to the formation of gaseous species. This drawback can be overcome by loading the liquid phase with an inert filler (such as a fine powder of the ceramic matrix) and/or by repeating several times ($n = 4$–10) the impregnation–curing–pyrolysis sequence, hence increasing processing time and labor cost at the plant level. Further, organometallic liquid precursors are often sensitive to moisture. The PIP process relies on relatively simple technologies.

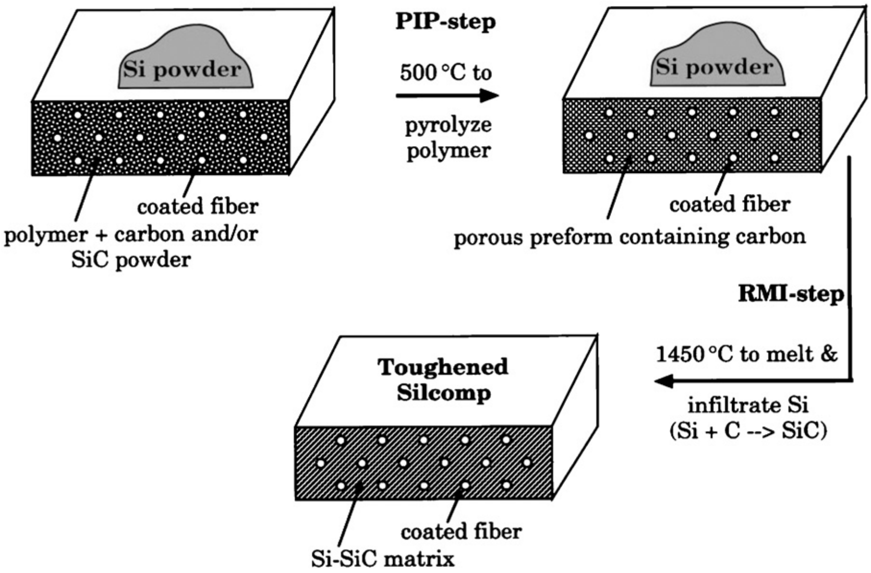


Figure 2
Principle of RMI processing of Si–SiC matrix composites (courtesy of K. Luthra).

(*c*) *Slurry impregnation and hot pressing* (*SIHP*). The SIHP process can be used to fabricate a variety of oxide and nonoxide matrix composites. In a first step, a fiber tow (or a fiber tape) is impregnated with a slurry, i.e., a stable suspension of matrix powder in a liquid containing a fugitive organic binder and various additives. The wet continuous material is taken up on a drum, yielding after drying a green composite prepreg somewhat similar to those employed for producing polymer matrix composites. After cutting and stacking, the prepreg layer assembly is heated to eliminate the fugitive binder and hot pressed yielding a dense matrix composite (Mazdiyasni 1990).

Two requirements should be fulfilled. First, the fibers should be stable under the hot-pressing conditions; hence, the fiber choice is limited to the most stable ceramic fibers, such as the oxygen-free SiC fibers. Second, the matrix should soften enough and flow during hot pressing to embed the fibers, supposing that some liquid phase is formed. This is typically the case for glass and glass–ceramic silica-based matrices, as already mentioned, the hot-pressing temperature being 1200–1400 °C. For such matrices, a heat treatment (the ceramization step) is usually performed after hot pressing, to transform partly or totally the glassy matrix into the related stable crystalline phase(s). Specific sintering additives, which form a small amount of liquid phase at the sintering temperature, have to be added to the slurry, for refractory covalent nonoxide matrices (such as SiC and Si_3N_4), hot pressing being performed at much higher temperatures (up to ~1800 °C), with some risk of fiber degradation. Some organic or organometallic liquid precursors (e.g., PCS for SiC) can also be added to the slurry. An important advantage of CMCs fabricated according to the SIHP process lies in their low residual porosity.

2.2 Vapor-phase Routes

In vapor-phase routes, also referred to as chemical vapor infiltration (CVI), the interphase and the matrix are successively deposited within the pore network of a fiber preform from gaseous precursors. The chemical reactions involved in the deposition process have to be activated. This is achieved by heating the fiber preform to moderate temperatures (typically 900–1100 °C) either in cold- or hot-wall reactors, the latter being usually preferred. CVI is closely related to chemical vapor deposition (CVD), a process commonly used in material engineering to form thin films, thick coatings, or filaments. Both processes have the same chemical basis but are run under different temperature (T), pressure (P), and gas flow rate (Q) conditions. In CVD, deposition mainly occurs on the external surface of the substrate, whereas in CVI it takes place within the pores of the substrate, the deposition process being used to embed the fibers in the ceramic matrix, i.e., to densify the fiber preforms (Warren 1992).

Two different phenomena are involved in CVI: (a) chemical reactions occurring either in the gas phase or on the fiber surface with formation of the deposit and (b) mass transfer of the gaseous reactants and products outside the preforms and in their pore networks. As an example, carbon, silicon carbide, boron nitride, and alumina can be deposited according to the following overall equations (Naslain and Langlais 1986):

$$CH_4(g) \rightarrow C(s) + 2H_2(g) \tag{1}$$

$$C_3H_8(g) \rightarrow 3C(s) + 4H_2(g) \tag{1a}$$

$$CH_3SiCl_3(g) \overset{H_2}{\rightarrow} SiC(s) + 3HCl(g) \tag{2}$$

$$BX_3(g) + NH_3(g) \rightarrow BN(s) + 3HX(g) \quad \text{with} \quad X = F, Cl \tag{3}$$

$$2AlCl_3(g) + 3CO_2(g) + 3H_2(g) \rightarrow Al_2O_3(s) + 3CO(g) + 6HCl(g) \tag{4}$$

Actually, the chemical mechanisms are much more complex and often involve intermediate gaseous species, the gas phase undergoing a maturation process as it diffuses in the pores. Mass transfer of the gaseous species occurs by convection and/or diffusion depending on T–P–Q conditions. Generally speaking, a key requirement in CVI is that the pores should remain open up to the end of the densification process.

(*a*) *Isothermal/isobaric CVI* (*I-CVI*). In I-CVI, the fiber preforms are set in the isothermal zone of a hot-wall deposition chamber (Fig. 3). The gaseous precursor flows outside the preforms by convection and inside the preforms by diffusion. In order to avoid an early sealing of the pores entrance by the deposit, deposition is performed at relatively low temperatures (to lower reaction rates) and under reduced pressures (to favor the diffusion of the gaseous species in the pores). Hence, the deposition yield and deposition rate are low. When necessary, the densification process is stopped and the preform surface machined in order to reopen the porosity or to enlarge the pore entrance (deposition occurring preferentially near the pore entrance) (Naslain and Langlais 1986, Savage 1993, Warren 1992).

(*b*) *Temperature/pressure gradient CVI* (*F-CVI*). The F-CVI process (the "F" signifying "forced") has been designed to increase the deposition rate and to shorten the densification time via the use of temperature and pressure gradients applied to the

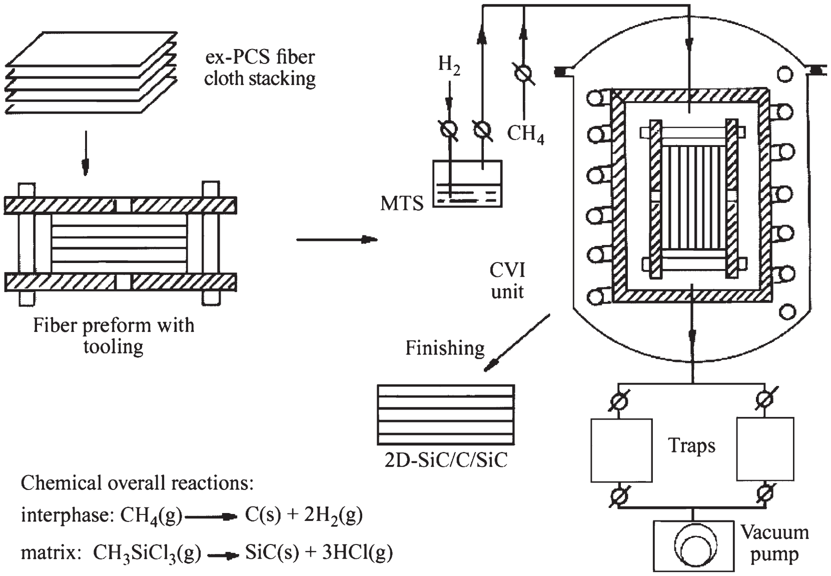

ex-PCS fiber cloth stacking
Fiber preform with tooling
H_2
CH_4
MTS
CVI unit
Finishing
2D-SiC/C/SiC
Traps
Vacuum pump
Chemical overall reactions:
interphase: $CH_4(g) \rightarrow C(s) + 2H_2(g)$
matrix: $CH_3SiCl_3(g) \rightarrow SiC(s) + 3HCl(g)$

Figure 3
Principle of I-CVI processing of 2D fiber/SiC matrix composites (courtesy of Woodhead Publishers).

fiber preform in opposite directions. The reactant species are injected in the fiber preform through its cold face (where almost no deposition occurs), under pressure, whereas the porosity is rapidly filled near the opposite hot face (where the deposition rate is high). As deposition proceeds, a densification front moves from the hot face to the cold face (whose temperature progressively increases if the deposit is conducting, this being the case for both pyrocarbon and SiC), while the unreacted precursor and reaction products are extracted by pumping first through the hot face of the sample then through its lateral face(s). Under such conditions, the densification time is significantly shortened (Mazdiyasni 1990).

(*c*) *Pressure-pulsed CVI* (*P-CVI*). In P-CVI, the furnace and the fiber preform pore network are first evacuated by pumping. Then the reactants are injected in a very short time in the reactor and inside the preforms. Deposition occurs during a given time, referred to as the residence time. This sequence of events is repeated in time, the opening and closing of the gas flow valves being operated automatically. One of the advantages of P-CVI lies in the fact that the fiber preform pore network is periodically fed with fresh reactants whereas the reaction products (which often act as reaction inhibitors) are periodically evacuated. Second, when the residence time is low enough (typically a few seconds), the deposit thickness per pulse can become extremely small (of the order of 1nm or less). Hence, if the composition of the precursor is periodically changed, P-CVI yields multilayered ceramic deposits. As an example, if one switches periodically from propane (Eqn. (1)) to methyltrichlorosilane/hydrogen (Eqn. (2)), a multilayered $(PyC–SiC)_n$ interphase can be deposited *in situ* on the fibers surface (Fig. 1(a)) (Naslain 1998).

(*d*) *Film boiling*. In the film boiling (or calefaction) process, a self-standing fiber preform, inductively heated with a susceptor, is directly immersed in a boiling liquid precursor, e.g., cyclohexane for pyrocarbon matrices. Assuming an axisymmetrical configuration, i.e., a cylindrical preform with a hole at the center in which the susceptor is set, a steep temperature gradient is generated radially in the preform. Further, precursor mass transfer in the preform, diphasic in character, is extremely efficient. As in the F-CVI process, a densification front moves vs. time from the preform hot surface (that near the susceptor) to the surface in contact with the boiling precursor. Densification times of the order of one day have been reported (David *et al.* 1999).

(*e*) *Advantages and drawbacks*. CVI processes are run at relatively low temperatures and hence are less damaging to the brittle fibers. They also yield interphases and matrices with high microstructural quality and excellent mechanical properties. Further, some of these processes are extremely flexible and all yield near-net-shape materials. However, CVI processes require important investments for volume production and the resulting CMC parts exhibit some residual open porosity. This residual porosity is not a problem from a mechanical standpoint since CMCs are tough materials; however, the occurrence of open residual porosity (that does not exist in composites fabricated according to some of the liquid-phase routes, e.g., the RMI process for SiC–Si matrix composites) favors the in-depth diffusion of oxygen in service conditions, which can be a drawback in nonoxide CMCs unless the materials are self-healing.

I-CVI has been already transferred to the plant level for the fabrication of C/C brake disks or C/SiC jet engine parts. Its main advantage is its high flexibility, a large number of preforms (which may have different sizes and shapes) able to be treated simultaneously. Even when intermediate surface machining of the partly densified preforms is performed, there is still much less handling than in, say, the PIP process. The flexibility of the process mitigates the fact that the densification times are relatively long. I-CVI is the preferred process for the volume production of CMC parts.

Both F-CVI and film boiling are attractive processes since densification times can be reduced by almost one order of magnitude with respect to I-CVI; however, their flexibility is much lower. In F-CVI, a specific fixturing has to be used for each fiber preform to establish the temperature/pressure gradients. As a result, this process is not suitable for the densification of preforms of complex shapes. In a similar manner, the use of a susceptor per preform (or pile of preforms) somewhat limits the flexibility of the film boiling process. To summarize, both processes may be advantageous when limited series of parts, with a relatively simple shape, have to be fabricated and delivered in a short time.

P-CVI is still an experimental process. It has a flexibility similar to that of conventional I-CVI but it requires more complex facilities. Although its use may result in shorter densification times, its main advantage lies in the fact that it is extremely suitable for the infiltration of highly engineered multilayered ceramics (interphases and matrices) in porous fiber preforms (Fig. 1).

2.3 Solid-state Routes

Solid-state routes, which are in principle closely related to the powder processing of multiphase ceramics, are limited to CMCs with short fiber reinforcements, such as whiskers (Mazdiyasni 1990). The starting material is a homogeneous powdered mixture of the fibers and the matrix, with appropriate additives (e.g., fugitive binder, sintering aids, stabilizers, etc.). It is often prepared by ball-milling in a protective liquid medium. After drying, the powder mixture is poured into the die of a press and sintered by uniaxial hot pressing. It can also be cold pressed, encapsulated in a suitable material (a metal or a glass), and sintered isostatically. During the ball-milling and hot-pressing steps some fragmentation of the brittle fibers usually occurs, leading to a decrease of the fiber aspect ratio and reinforcing capability. Further, the presence of the short fibers often renders the sintering of the powder matrix still more difficult (sintering of covalent carbides and nitrides, such as SiC and Si_3N_4, being known to be intrinsically difficult and to necessitate sintering additives), requiring high temperatures and pressures. Hence, only very stable fibers can be used (short carbon fibers or SiC whiskers). Finally, composites fabricated with short fibers according to solid-state routes display mechanical properties which are much lower than those of the related composites processed from continuous fibers.

SiC whiskers have been used in alumina-based or silicon nitride matrices, e.g., for producing cutting-tool inserts and other wear-resistant items, with an increase in toughness, hardness, and lifetime. It is important to note that the use of whiskers of small size raises health problems somewhat similar to those reported for asbestos fibers, which actually limit the development of these composites.

3. *Concluding Remarks*

The development of low-cost, high-performance CMCs relies on available fibers and it supposes an optimization of both the FM interfacial zone and the matrix itself. Some progress has been achieved in the field of ceramic fibers. As an example, oxygen-free SiC fibers and quasistoichiometric SiC fibers display a higher thermal stability than their Si–C–O counterparts. Hence, they are compatible with a variety of processing techniques including the high-temperature processes. However, available oxide fibers are not stable enough from a microstructural standpoint.

A key point in terms of material design is to improve the durability of nonoxide CMCs in oxidizing atmospheres. Taking the example of SiC/SiC composites, a promising approach is to replace the pyrocarbon interphase by boron nitride or multilayered $(PyC–SiC)_n$ or $(BN\text{-}SiC)_n$ interphases and the SiC matrix by self-healing multilayered matrices. In terms of processing, rapid densification techniques (such as RMI or film boiling) are becoming available beside well-established routes (such as CVI) which could still be improved. Even if these processes display less flexibility, they may be appropriate for specific matrix compositions and/or specific application areas.

See also: Carbon Fibers*; Glass Fibers*; Whiskers*

Bibliography

Bunsell A R, Berger M H 1999 *Ceramic Fibers*. Dekker, New York

Chawla K K 1993 *Ceramic Matrix Composites*. Chapman & Hall, London

David P, Benazet J D, Ravel F, 199 Rapid deposition of carbon and ceramic matrices by film boiling technique. In: Vincenzini P (ed.) *Advanced Structural Fiber Composites*. Techna, Faenza, Italy, pp.135–40

Fitzer E, Manocha L M 1998 *Carbon Reinforcements and Carbon/Carbon Composites*. Springer, Berlin

Lamouroux F, Bertrand S, Pailler R, Naslain R, Cataldi M 1999 Oxidation-resistant carbon-fiber-reinforced ceramic-matrix composites. *Compos. Sci. Technol.* **59**, 1073–85

Luthra K L, Singh R N, Brun M K 1993 Toughened Silcomp composites: process and preliminary properties. *Am. Ceram. Soc. Bull.* **72** (7), 79–85

Mazdiyasni K S 1990 *Fiber Reinforced Ceramic Composites*. Noyes, Park Ridge, NJ

Naslain R R 1998 The design of the fibre–matrix interfacial zone in ceramic matrix composites. *Composites*. **29A**, 1145–55

Naslain R R 1999 Materials design and processing of high temperature ceramic matrix composites: state of the art and future trends. *Adv. Compos. Mater.* **8** (1), 3–16

Naslain R R, Langlais F 1986 CVD processing of ceramic–ceramic composite materials. *Mater. Sci. Res.* **20**, 145–64

Savage G 1993 *Carbon–Carbon Composites*. Chapman & Hall, London

Urquhart A W 1991 Novel reinforced ceramics and metals: a review of Lanxide composite technologies. *Mater. Sci. Eng. A* **144**, 75–81

Warren R 1992 *Ceramic Matrix Composites*. Blackie, Glasgow, UK

R. R. Naslain
Université Bordeaux 1, Pessac, France

Ceramic Nanocomposites with Organic Phases, Optics of

Since the 1970s there has been considerable progress in the sol–gel technique for the manufacture of glasses, glass-ceramics and ceramics. This method has been successfully used for a variety of products ranging from bulk glasses (Kirkbir *et al.* 1996) to optical fibers (Susa *et al.* 1982), to special coatings, ultra-pure powders and multifunctional materials. Also, the sol–gel processed transparent porous matrix offers an exciting potential as a host matrix for doping optically active organic molecules (Nogues and Moreshead 1990). This matrix exhibits many properties of inorganic glasses as well as some unique characteristics that can lead to novel developments in photonics. Avnir *et al.* (1984, 1985) demonstrated for the first time that organic molecules can be added to the sol–gel matrix and thin films. Since then, many groups around the world have introduced organic materials into a variety of inorganic matrices via the sol–gel process. Sol–gel processed materials have been used in the areas of solid-state lasers and in platforms for chemical and biosensors (Braun *et al.* 1990, Narang *et al.* 1993). The sol–gel technique has also been used to fabricate nonlinear optically active composites for applications in optical telecommunications (Burzynski and Prasad 1994). Organically modified and sol–gel processed materials were shown to exhibit excellent nonlinear optical χ^2 and χ^3 properties. The sol–gel method also provides a more convenient route to prepare luminescent glass optical fibers doped with rare earth ions (Wu *et al.* 1993, 1994), waveguides (Bahtat *et al.* 1994, Benatsou *et al.* 1997), and lasers (Thomas *et al.* 1992).

To date, there are only a few sol–gel-derived materials that meet the materials quality required for fabricating devices. One of the major problems has been to produce useful bulk materials with controlled doping (organic and inorganic) at the nanoscale. Another problem has been the fabrication of a low-loss optical fiber with active species incorporated within it.

This article briefly discusses different sol–gel approaches, which have shown promising results in producing useful materials for photonics. They include (i) rare earth doped glasses, (ii) multiphasic nanostructured composites, which combine the merits of the inorganic glass and an organic polymer, and (iii) optical fibers with active molecules incorporated within the fiber matrix and dispersed homogeneously for sensing or lasing applications. Selected examples are given of the preparation processes and applications based on the past and current results obtained primarily in the authors' laboratory (see also *Reactive Ceramic Nanocomposites with Organic and Bio-organic Phases**).

1. *Sol–Gel Processed Materials*

1.1 Sol–Gel Process

Sol–gel processing involves the hydrolysis of a metal alkoxide, followed by a cascade of condensation and polycondensation reactions. More detailed information pertaining to the chemistry of sol–gel processing can be found in several books and review articles (Brinker and Scherer 1990, Hench and West 1990).

A sol is a dispersed system where the dispersed phase, consisting either of solids, discrete "globules" or oligomers in the form of chains, is present in a liquid dispersion medium. The medium can be water or any other liquid. A gel is a stiff mass of a continuous phase, more like a viscous liquid. The sol can be prepared by controlled hydrolysis and polycondensation reactions of alkoxides, which then form the network of the resultant glasses.

The most commonly used precursor materials for the silica sol formation are tetramethylorthosilicate (TMOS) and tetraethylorthosilicate (TEOS). The basic reactions of a sol–gel system undergoing concurrent hydrolysis and condensation are:

$$n\mathrm{Si(OR)_4} + 4n\mathrm{H_2O} \rightarrow n\mathrm{Si(OH)_4} + 4n\mathrm{ROH} \qquad (1)$$

$$n\mathrm{Si(OH)_4} \Rightarrow n\mathrm{SiO_2} + 2n\mathrm{H_2O} \qquad (2)$$

Gelation occurs when the growing polymers cross-link to form an extensive three-dimensional network. Because of the sharp increase in viscosity accompanying gelation, it may be considered a rapid solidification process. In other words, its structure at a given state of polymerization is essentially "frozen in" at the gel point. Aging under conditions in which appreciable polymer solubility occurs (e.g., high water and high pH) can result in extensive restructuring of the gel. The strength of the wet gel increases with aging, which reduces the chance of cracking during drying.

A promising class of sol–gel-derived materials is the organic–inorganic hybrid materials which combine the merits of an inorganic glass and an organic polymer. A novel way to prepare high optical quality nanocomposite monoliths with desirable mechanical behavior is by forming a rigid glass monolith matrix with controlled porosity through the sol–gel process. Then, an organic monomer is diffused into the open pores and polymerized inside. Pope *et al.* (1989) pioneered this method.

1.2 Sol–Gel Silica Glasses and Composites

The silica skeleton of a composite glass is prepared by hydrolyzing silicon ethoxides with an acid or a base followed by heat treatment. The resulting product is a highly porous monolith with a wide range of pores that can be controlled by manipulating different process parameters. The porous monoliths are either densified thermally to make silica glasses or they can be impregnated with an organic phase resulting in composite glasses. More detailed preparation procedures for silica and composite glasses are explained below.

The silica glasses doped with rare earth ions were prepared according to the procedure described elsewhere (Biswas *et al.* 1999). Briefly, TEOS was hydrolyzed with a 0.1 MNH_4OH solution. The molar ratio of TEOS:H_2O was 1:10. The mixture was magnetically stirred to obtain a clear homogenous solution (sol). The sol was cast into a polystyrene Petri dish and covered with a thin polyethylene film. The sol transformed into a gel within 24 hours at room temperature. The gels were aged and dried at 45 °C for 2 weeks and finally heat treated at 1000 °C to form a rigid porous silica matrix with a pore volume of 60% and an average pore size of 100 Å (measured using nitrogen adsorption isotherms and calculated by the BET method).

The porous silica gels were soaked with a different dopant salt solution for 12 hours. Simple chemical treatments (Maeda *et al.* 1990) were used to immobilize the dopant ions in the pores to ensure uniform ion distribution in the bulk of the final glasses. After removal from the solutions, the gels were dried at 50 °C for 12 hours and then heat treated at ~1150 °C for 1 hour to convert into dense pore and bubble free glasses.

The fabrication of the composite glass was done following the method described by Gvishi *et al.* (1995a, 1997). Highly porous silica-gel bulk glasses were prepared by a two-step hydrolysis sol–gel process (Pope *et al.* 1989). The procedure starts with a precursor solution, which contained TEOS and ethanol in the molar ratio 1:4. After one hour of pre-hydrolysis, water (molar ratio 4) and HNO_3 (molar ratio 0.06) were added to complete the hydrolysis. This step was followed by the addition of HF (molar ratio 0.8) to achieve fast gelation. The solution was then covered with a lid having small orifices to allow the solvents to escape and was kept for aging at 40 °C for 1–3 weeks. The bulk gels were then partially densified by slow heating (60 °C/hour) from room temperature to 500 °C. These gels had an initial surface area of about 850 $m^2\,g^{-1}$, a pore volume of 68%, and an average pore size of 46 Å. These bulk glasses (dimensions 10 mm × 5 mm × 1–5 mm) were immersed in a methylmethacrylate (MMA) monomer solution for 10–15 min. The MMA-dye solution diffused into the sol–gel-derived glass pores and was polymerized therein using benzoyl peroxide (2%) as the initiator. The MMA-dye-doped bulks were re-immersed in a MMA-dye solution, which at this stage was initiated for full polymerization with benzoyl peroxide (0.5%), and kept in a sealed container at 40 °C until the polymerization process was completed (a few days). After the completion of the polymerization process, the samples were withdrawn, cleaned with chloroform, and polished to obtain parallel-piped slabs with clear smooth surfaces.

1.3 Sol–Gel-derived Fibers

The process of drawing continuous fibers using the sol–gel technique has been studied extensively. Under specific preparation conditions, fibers up to 1 m in length can be pulled from the sol by hand or using a spinning apparatus.

A stock sol–gel solution was prepared by stepwise mixing of tetraethoxysilane, water, ethanol, and hydrochloric acid in molar ratio of 1:1.5:1.9:0.01, respectively. The ethanol solution used in the above step contained the dye to be entrapped within the sol–gel matrix. The dye-doped sol–gel solution was placed in a beaker, sealed with aluminum foil, and stored in an oven at 40 °C for gelation (several days). Just before gelation (indicated by the high viscosity of the sol), sol–gel-derived uniform continuous fibers were drawn. The fibers were kept for aging and drying under ambient conditions for a few weeks. They were cut to a specific length without any polishing of the fiber ends. The average attenuation loss was 60 dBm^{-1} from the organically doped sol–gel-derived fiber. However, in each fiber there are several sections with losses as low as 10 dB m^{-1} (Narang *et al.* 1996).

The potential of the dye-doped sol–gel-derived optical fibers has been studied by doping with fluorescein, a well-known pH sensitive fluorescent probe (Narang *et al.* 1996) for chemical fiber sensing and with rhodamine-6G, a well-known laser dye, as a fiber laser medium (MacCraith *et al.* 1991, Drexhage 1990).

2. Applications of Sol–Gel Glasses, Composites, and Fibers

2.1 Sol–Gel Processed Rare Earth Doped Glasses

Silica glasses doped with rare earth ions are potential materials for solid-state fiber lasers and amplifiers, memory devices, switches, sensors, etc. Early work on silica glasses reveals that silica can be doped with rare earth ions to a maximum concentration of 0.1 mol.% by a conventional melt technique.

Interest in high-powered laser glasses of rare earth ions in a silica matrix, due to its favorable thermal and mechanical properties, has prompted several researchers to increase the solubility of rare earth ions in silica glasses through various techniques. Pope and Mackenzie (1988) first suggested that the sol–gel method could be effectively used to obtain relatively large concentrations of neodymium in a silica host as compared to that obtained by conventional melting or plasma torch CVD techniques. The works of others (Moreshead *et al.* 1990) have focused on the development of bulk neodymium-doped sol–gel silica glasses for laser applications. Among the other rare earth ions, Er^{3+} (Zhou *et al.* 1997), and Eu^{3+} (Nogami and Abe 1996) have been studied most extensively. However, their success was severely restricted by the fact that the rare earth ions tend to have a nonuniform distribution and form clusters (Lee *et al.* 1994) in the bulk, which reduces the fluorescence efficiency and lifetime. The presence of a large number of hydroxyl groups (Fujiyama *et al.* 1991) in gel glasses also degrades the emission properties of the rare earth ions. In order to prepare device-quality silica glasses doped with rare earth ions, hydroxyl quenching and ion clustering must be eliminated. Nevertheless, laser emission has been observed in a neodymium-doped sol–gel glass (Thomas *et al.* 1992).

It was shown that the Al^{3+} co-doping prevents the clustering of rare earth ions (for the atomic ratio of Al^{3+} to Nd^{3+} greater than 10) and disperses them homogeneously in the SiO_2 glass matrix (Arai *et al.* 1986). Therefore, it overcomes the concentration quenching problem in sol–gel-derived rare earth doped glasses. However, the addition of alumina was found to promote hydroxyl quenching by retaining more hydroxyl groups in the glass (Lochhead and Bray 1995). Several other methods have been proposed to reduce rare earth ion clustering, including using acetate precursors (Chakrabarti *et al.* 1994), ion encapsulation by chelating ligands and co-doping with other metallic ions (Costa *et al.* 1996).

Hydroxyl-free silica glasses by the sol–gel method can be realized by treating the gels with either CCl_4 (Phalippou *et al.* 1984) or Cl_2 (Satoh *et al.* 1997) at high temperature. However, this process is not suitable for rare earth doped gels because of the volatile nature of the chloride byproduct. Another proposed method to reduce the hydroxyl group is by doping the gel with an index lowering element, such as fluorine (Rabinovich *et al.* 1986).

Rare earth (RE) ions are generally incorporated into a glass in their trivalent state. However, some of the RE ions can exist in a divalent state as well. These divalent RE ions show very strong emission because of the dipole-allowed electronic transition. Sm^{2+} and Eu^{2+} doped glasses have attracted a great deal of attention for applications in memory devices and as phosphor materials. The broad emission of Eu^{2+} ions is host sensitive. Depending on the host matrix, Eu^{2+} ions could emit from the UV to the red region of the electromagnetic spectrum.

Divalent rare earth ion doped porous gel glasses can be prepared by heating the gel in the presence of a H_2–N_2 gas mixture (Nogami *et al.* 1998) at a relatively high temperature. The Eu^{3+} ions were also shown to be readily reduced to Eu^{2+} in a silica xerogel at room temperature by a "surface assisted reduction" (Zaitoun *et al.* 1998). In another approach, *in situ* formation of hydrogen in a mixed zirconium propoxide and siloxane sol (Cordoncillo *et al.* 1998) was found to reduce the Eu^{3+} ions at room temperature.

Recently, it was shown that the Eu^{3+} ions were spontaneously reduced in the sol–gel matrix during densification (Biswas *et al.* 1999). The optical absorption spectrum in the 200–500 nm range of the Eu–Al co-doped glasses densified at 1125 °C in air is shown in Fig. 1. The spectrum shows a broad peak around 266 nm that is attributed to a transition of the Eu^{2+} ions from $4f^7$ to $4f^65d$ (t_g). The emission and the excitation spectra of the Eu–Al co-doped sample at 1125 °C are shown in Fig. 2. The emission shows a broad intense peak at 435 nm. A qualitative measurement based on the emission intensity and the absorption coefficient of these glasses shows that a maximum 85% of the Eu^{3+} ions were converted to Eu^{2+} ions. Figure 3 shows the excitation spectra for the $^5D_0 \rightarrow {}^7F_2$ transition of Eu^{3+} ions (614 nm) in the densified silica and in the Al^{3+} co-doped silica matrices. The sharp lines observed between 200 and 500 nm are assigned to the *f–f* transitions of the Eu^{3+} ions in the glasses. However, a broad band with a peak at 300 nm is observed only in the aluminium co-doped sample that contains both Eu^{2+} and Eu^{3+} ions. The excited Eu^{2+} ions in the $4f^65d$ (e_g) level can

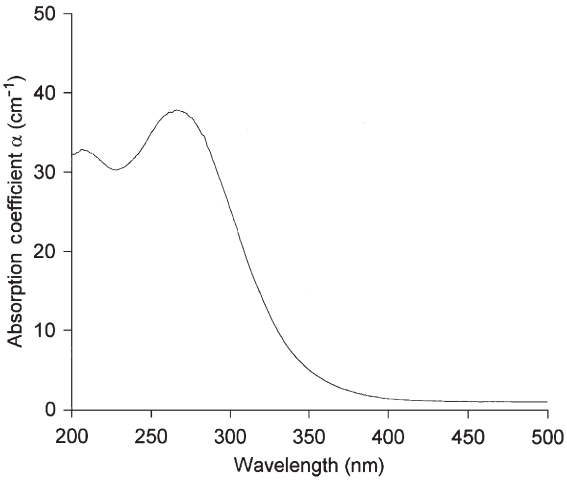


Figure 1
The absorption spectrum of a Eu–Al co-doped silica glass.

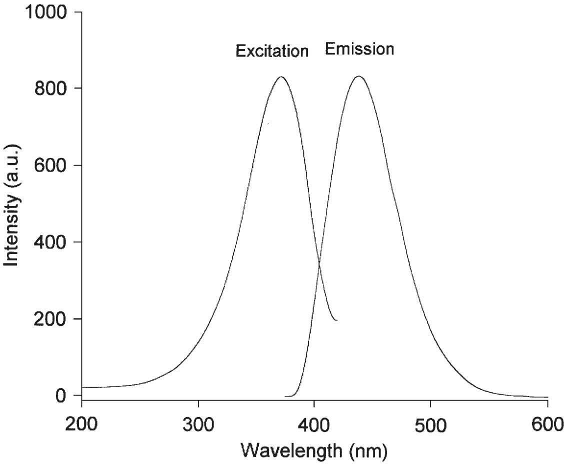


Figure 2
The excitation and the emission spectra of a Eu–Al co-doped silica glass.

relax to the ground state by transferring energy to the Eu^{3+} ions which, in turn, are excited to the $^5D_2$ level.

2.2 Multiphasic Nanostructured Composites

Sol–gel derived composite glasses, due to their excellent optical quality, play an important role in photonics applications such as lasing, optical power limiting, nonlinear optical response, etc. (He *et al.* 1995, Gvishi *et al.* 1995a, 1995b). The multiple impregnation method can be used to dope two (or more) different optically active materials, with each existing in different phases of the matrix (the inorganic silica phase, the organic polymer phase, and the interfacial phase).

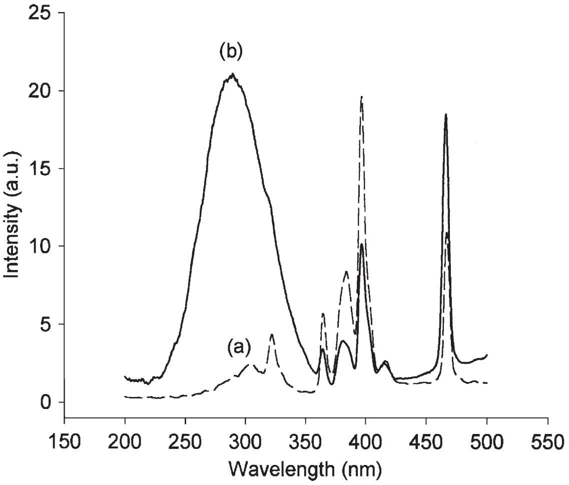


Figure 3
The excitation spectra for the $^5D_0 \rightarrow {}^7F_2$ transition of Eu^{3+} ions in (a) SiO_2 and (b) Al_2O_3–SiO_2 glasses.

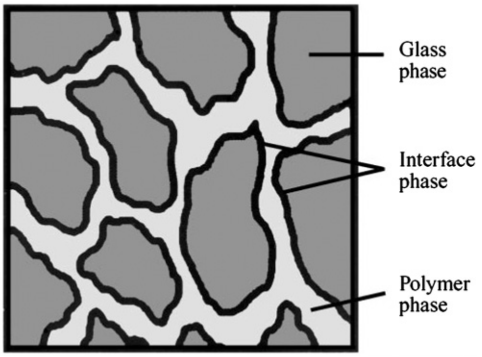


Figure 4
A schematic representation of the multiphasic composite glass.

The basic idea of the multiphasic nanostructured composite is illustrated in Fig. 4. The inorganic phase (phase #1) is doped with inorganic species that can survive the heat treatment. The pores can be doped with organic functional molecules through impregnation of the solution, or by using a melt or vapor-phase infiltration. After draining out the solvent and drying, the dopant organic molecules form a phase (phase #2) on the surface of the pores. Finally, the pores are filled with an index-matched polymer by *in situ* polymerization of a monomer, such as methylmethacrylate. Another active organic species can also be introduced in the polymer phase (phase #3). The resultant monoliths, thus prepared, are suitable for many applications in photonics.

(*a*) *Solid-state dye laser*. Laser emission via a one-photon excitation process (He *et al.* 1994) or a two-photon excitation process (Zhao *et al.* 1995, He *et al.* 1996) from dye doped solid matrices has been reported. These composite glasses exhibit excellent slope efficiencies (up to 39%). The success of a solid-state dye laser depends mainly on its photostability and mechanical properties. Sol–gel derived composite glasses offer better photostability, optical quality, and mechanical stability as compared to other organic matrices. Solid-state dye lasers could be useful for applications where only low laser output is required, e.g., photodynamic therapy (PDT), medical diagnostics of tissues using laser-induced fluorescence (LIF), and photodynamic purification of blood from viruses.

The sections below present the results obtained in our laboratory on (i) multiexcitation pumping (one- and two-photon excitation) of a dye-doped composite glass and (ii) multidye laser—a multiphasic nanostructure composite glass monolith containing two different laser dyes.

(*b*) *One- and two-photon pumped dye laser*. Sol–gel composite glasses doped with ASPD in the interfacial surfaces have been prepared according to the procedure discussed above. Lasing experiments were carried out by one-photon excitation pumping. The laser output spectrum of the composite glass doped with ASPI ($\sim 3.1 \times 10^{-3}$ M), pumped with a frequency-doubled Nd:YAG laser is shown in Fig. 5. The output exhibits a peak at ~593 nm with a full width at half maxima (FWHM) of 2 nm (Zhao *et al.* 1996).

Figure 6 shows the output energy vs. input pump energy for ASPD in the composite glass (Zhao *et al.* 1996) pumped with a frequency-doubled Nd:YAG laser (532 nm). The observed threshold was 0.244 mJ/pulse for the composite glass doped with $\sim 3.1 \times 10^{-3}$ M ASPI. Maximum lasing slope efficiencies of ~9% in the composite glass was observed under one-photon excitation. From calculations, it was found that there was no evidence for additional losses, such as excited-state absorption, which were observed for red perylimide dye-doped composite glass (Gvishi *et al.* 1993).

ASPI also exhibits a strong two-photon absorption at 1.06 μm followed by red fluorescence. Using an appropriate optical cavity and longitudinal pumping with an intense IR beam, up-converted visible lasing was obtained from ASPI-doped composite glasses. Two-photon lasing was first demonstrated in a polymer matrix (He *et al.* 1996) and subsequently in composite glass and composite vycor glass (Zhao *et al.* 1995). The lasing efficiency was found to be very high (~3.5%) as opposed to the expectation based on the assumption that two-photon absorption is not an efficient process. For a dye-doped polymer, the lasing lifetime in terms of pulse number was more than 4×10^4 pulses at a 2 Hz repetition rate and with a pump energy of 1.3 mJ. Since, it has been demonstrated that the photostability of a sol–gel composite glass matrix is much higher than those of a polymer host (three orders of magnitude) under one-photon excitation; one can expect the same under two-photon pumping. It was found that the composite glasses could stand a flux power of 400 MW cm^{-2} with 10 ns pulses without damage (Zhao *et al.* 1995, He *et al.* 1996).

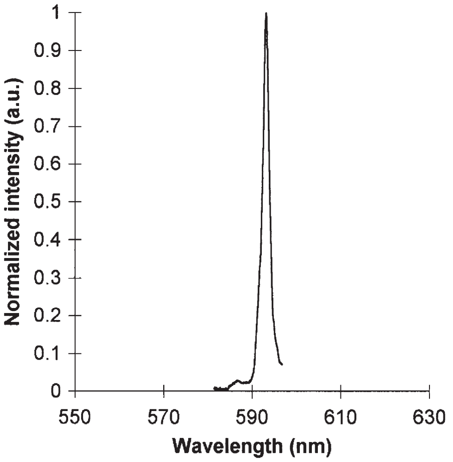


Figure 5
Lasing output vs. wavelength for the ASPD composite glass. The composite glass was transverse pumped with an 8 ns pulsed-frequency-doubled Nd:YAG laser at 532 nm and at a 1 Hz repetition rate.

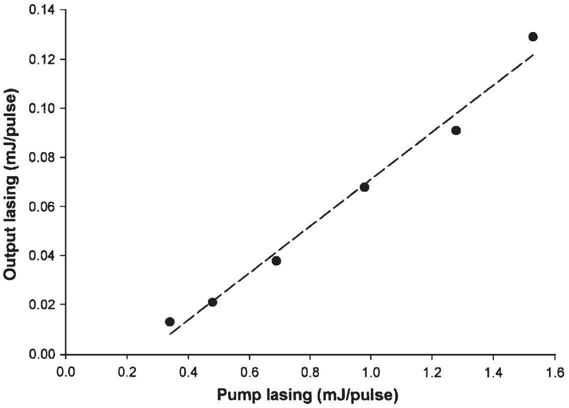


Figure 6
The output energy vs. input energy, under excitation with a frequency-doubled Nd:YAG laser (532 nm) for ASPD in composite glass.

(*c*) *Multidye laser*. The multidye solid-state composite glasses for tunable laser operation contain ASPI in the interfacial phase and R6G in the polymer phase of the matrix. The details of the fabrication process of the glasses is reported elsewhere (Ruland *et al.* 1996). Lasing was observed across the region of the individual dyes.

Figure 7 presents the lasing tunability of a multiphasic composite glass containing both dyes. Lasing tunability of the two composite glasses containing each dye separately is also shown in the same figure for a reference. The FWHM of the tunability spectra for the ASPI composite glass is ~21 nm and that for R6G is ~12 nm. However, a composite glass containing both dyes is ~37 nm. From this data it is evident that the laser emission from the composite glass containing both dyes is tunable across the range of both dyes (560–610 nm), whereas in the solution state the R6G emission is quenched. It was assumed that the quenching in the solution state is a result of Förster energy transfer. Therefore, due to the extremely large ratio between the specific surface area and the pore volume, which is ~8.5×10^6, it is possible to separate the two dye molecules (which reside in different phases) to a distance where the Förster energy transfer is no longer effective. This result illustrates the fascinating possibilities embodied in the multiphasic composite glasses for fabricating multifunctional devices for photonics.

2.3 *Organically Doped Fibers*

(*a*) *Micron-scale chemical and biosensor fibers*. Micrometer- and submicrometer-scale fiber optic sensing was first demonstrated by Tan *et al.* (1992). Here, micrometer-scale chemical and biosensor fiber applications are presented using a fluorescein-doped sol–gel-derived fiber, where the fluorescein is doped in the sol–gel matrix. Fluorescein has previously been doped within the sol–gel-derived monoliths and thin films as a pH sensitive dye (Reisfeld *et al.* 1987, MacCraith *et al.* 1991). The fluorescence quantum yield of fluorescein is pH dependent; increasing with increasing pH. The effect of pH on the fluorescence intensity by exposing the fluorescein-doped sol–gel-derived fiber (20 ± 3 μm) to ammonia and HCl vapors was investigated in our laboratory. The bulk of the instrumentation and the setup of the measurements were described before.

Figure 8 presents the change (increase) in fluorescence intensity of the fluorescein-doped sol–gel-derived fiber by exposing the fiber to ammonia and HCl vapors (Narang *et al.* 1996). A 50% increase in the fluorescence intensity was observed when it was exposed to ammonia vapors. A 45% decrease in the

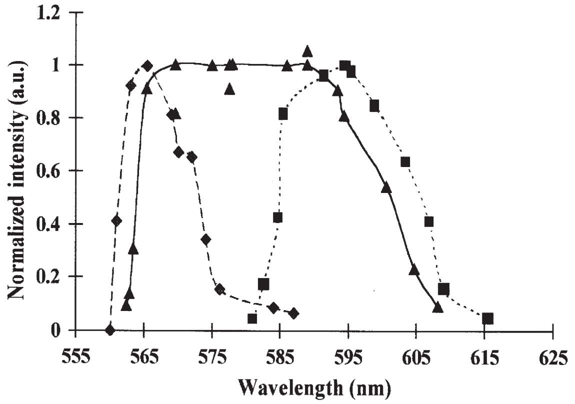


Figure 7
Lasing tunability of Rhodamine-6G composite glass (◆), ASPD composite glass (■), and the multiphasic composite glass containing both Rhodamine-6G and ASPD (▲). The composite glasses were transversely pumped with an 8 ns pulsed frequency-doubled Nd:YAG laser at 532 nm operating at a 1 Hz repetition rate, using a grating as the back reflector.

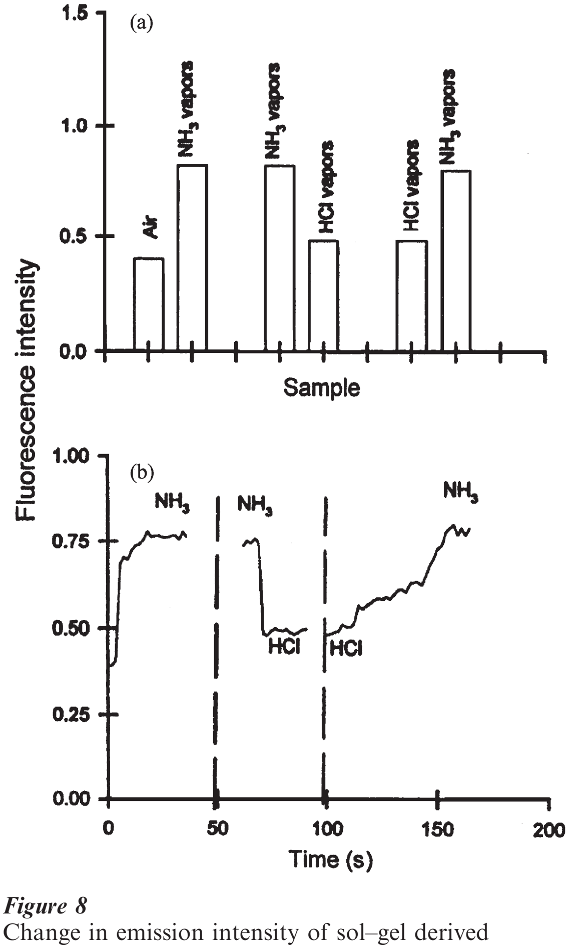


Figure 8
Change in emission intensity of sol–gel derived fluorescein-doped fiber under HCl or NH_3 exposure.

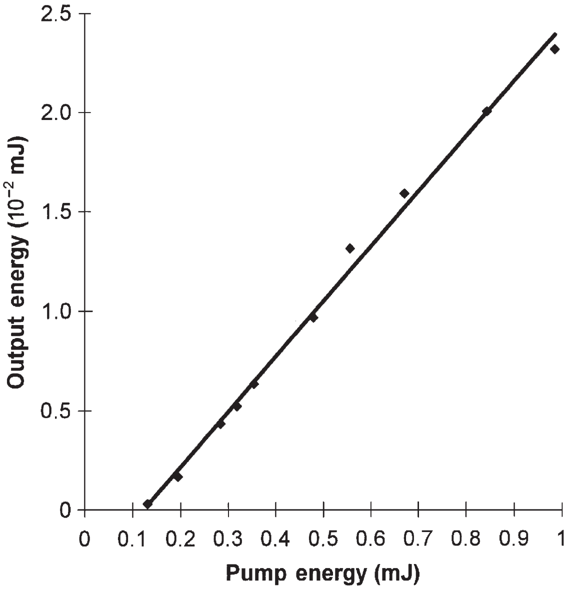


Figure 9
Slope efficiency of sol–gel derived Rodamine-6G doped fiber (length = 1 cm, diameter = 150 μm) laser pumped at 532 nm with a 8 ns pulsed frequency-doubled Nd:YAG laser.

fluorescence intensity was noted when the same fiber was treated with HCl vapors. The results indicate that a fluorescein-doped sol–gel-derived fiber could be used as a pH sensor. The response time of the fluorescein-doped sol–gel-derived fiber to ammonia and HCl vapors was observed to be 1–2 s (panel B), which was increased by an order of magnitude by further recycling. This method can be used to fabricate optical fiber based micron-scale sensors to detect *in situ* environmental changes. An improved response time is expected by using a smaller diameter fiber, as was achieved in the case of a pyranine-doped sol–gel-derived micropipette by Samuel *et al.* (1994).

(*b*) *Solid-state dye-doped fiber laser.* A solid-state fiber laser was demonstrated in our laboratory using R6G-doped sol–gel-derived fiber. R6G, a commercially available laser dye, exhibits very high lasing efficiency and solubility properties but low photostability.

An amplification of 10 was achieved in the cavity lasing compared to a single-pass super-radiance intensity (amplified emission under the conditions of a blocked back mirror). The lasing intensity was extremely sensitive to the alignment of the cavity component, due to the delicate coupling requirement into a micron-size fiber. Therefore, thick fibers (5.5 cm or 1 cm long, 150 μm in thickness) (Gvishi *et al.* 1996) were used for laser demonstration. Figure 9 presents the measured laser output–pump input intensity relationship for R6G ($\sim 10^{-4}$ M) doped sol–gel fiber (1 cm long, 150 μm in thickness), when pumped with an 8 ns pulse at a 1 Hz repetition rate (Gvishi *et al.* 1996). The observed lasing slope efficiency was ~2.8%. A much higher slope efficiency can be achieved by improving coupling into the fiber (by polishing the fiber ends), decreasing fiber optical loss, optimizing concentration-length relationship (preventing aggregation effect), and improving the cavity design.

With appropriate improvements in preparation and laser design, this device can be fabricated and operated as a single-mode dye-doped fiber laser. By replacing the back mirror with a grating, tunable laser action is expected.

3. *Summary*

Over the years, rapid progress has been observed in the field of sol–gel-derived materials for photonics. The present article focuses on specific applications of the sol–gel process, with emphasis mainly on rare earth doped glasses, a multidye solid state tunable laser, micrometer-scale chemical and biosensor fibers, and a solid-state dye-doped fiber laser.

Bibliography

Arai K, Namikawa H, Kumata K, Honda T, Ishii Y, Handa T 1986 Aluminum or phosphorus co-doping effect on the fluorescence and structural properties of neodymium-doped silica glass. *J. Appl. Phys.* **59**, 3430–6

Avnir D, Kaufman V R, Reisfeld R 1985 Organic fluorescent dyes trapped in silica and silica–titania thin films by the sol–gel method. Photophysical, film and cage properties. *J. Non-Cryst. Solids* **74**, 395–406

Avnir D, Levy D, Reisfeld R 1984 The nature of the silica cage as reflected by spectral changes and enhanced photo stability of trapped rodamine 6G. *J. Phys. Chem.* **88**, 5956–9

Bahtat A, Bouazaoui M, Bahtat M, Mugnier J 1994 Fluorescence of Er^{3+} ions in TiO_2 planar waveguides prepared by a sol–gel process. *Opt. Commun.* **111**, 55–60

Benatsou M, Capoen B, Bouazaoui M, Tchana W, Vilcot J P 1997 Preparation and characterization of sol gel derived Er^{3+}: Al_2O_3-SiO_2 planar waveguides. *Appl. Phys. Lett.* **71**, 428–30

Biswas A, Friend C S, Prasad P N 1999 Spontaneous reduction of Eu^{3+} ion in Al Co-doped sol–gel silica matrix during densification. *Mater. Lett.* **39**, 227–31

Braun S, Rappoport S, Zusman R, Avnir D, Ottolenghi M 1990 Biochemically active sol–gel glasses: the trapping of enzymes. *Mater. Lett.* **10**, 1–5

Brinker C J, Scherer G W 1990 *Sol–Gel Science. The Physics and Chemistry of Sol–Gel Processing*. Academic Press, Boston, MA

Burzynski R, Prasad P N 1994 New photonic media prepared by sol–gel process. In: Klein L C (ed.) *Photonics and Nonlinear Optics with Sol–Gel Processed Inorganic Glass: Organic Polymer Composite*. Kluwer, Boston, MA, Chap. 19

Chakrabarti S, Sahu J, Chakraborty M, Acharya H N 1994 Monophasic silica glasses with large neodymia concentration. *J. Non-Cryst. Solids* **180**, 96–101

Cordoncillo E, Viana B, Escribano P, Sanchez C 1998 Room temperature synthesis of hybrid organic–inorganic nanocomposites containing Eu^{2+}. *J. Mater. Chem.* **8**, 507–9

Costa V C, Lochhead M J, Bray K L 1996 Fluorescence line-narrowing study of Eu^{3+}-doped sol–gel silica: effect of modifying cations on the clustering of Eu^{3+}. *Chem. Mater.* **8**, 783–90

Drexhage K H 1990 Structure and properties of laser dyes. In: Schäfer F P (ed.) Dye Lasers. Springer, Berlin, Chap. 5, pp. 155–85

Fujiyama T, Yokoyama T, Hori M, Sasaki M 1991 Silica glass doped with Nd and Al prepared by sol–gel method: change in the state of aluminum in the formation process. *J. Non-Cryst. Solids* **135**, 198–203

Gvishi R, He G S, Prasad P N, Narang U, Li M, Bright F V, Reinhardt B A, Bhatt J C, Dillard A G 1995aa Spectroscopic studies of new blue laser dyes in tetrahydrofuran solution and in composite glass. *Appl. Spectrosc.* **49**, 834–9

Gvishi R, Narang U, Bright F V, Prasad P N 1995bb Probing the microenvironment of polymer-impregnated composite glass using solvatochromic dye. *Chem. Mater.* **7**, 1703–8

Gvishi R, Prasad P N, Reinhardt B A, Bhatt J C 1997 Third-order optical nonlinearity studies of *p*-heptaphenyl derivatives—doped sol–gel processed composite glass and THF solution by degenerate Four Wave Mixing and optical Kerr Gate measurements. *J. Sol–Gel. Sci. Tech.* **9**, 157–67

Gvishi R, Reisfeld R, Burshtein Z, Miron E 1993 New stable tunable solid-state dye laser in the red. *Proc. SPIE* **1972**, 390–9

Gvishi R, Ruland G, Prasad P N 1996 New laser medium: dye-doped sol–gel fiber. *Opt. Commun.* **126**, 66–72

He G S, Bhawalkar J D, Zhao C F, Prasad P N 1996 Properties of two-photon pumped cavity lasing in novel dye doped solid matrices. *IEEE J. Quant. Electron* **32**, 749–55

He G S, Gvishi R, Prasad P N, Reinhardt B A 1995 Two-photon absorption based optical limiting and stabilization in organic molecule-doped solid materials. *Opt. Commun.* **117**, 133–6

He G S, Zhao C F, Park C K, Prasad P N, Burzynski R 1994 Dye film leaky waveguide laser. *Opt. Commun.* **111**, 82–5

Hench L L, West J K 1990 The sol–gel process. *Chem. Rev.* **90**, 33–72

Kirkbir F, Murata H, Mayers D, Chaudhari S R, Sarkar A 1996 Drying and sintering of sol–gel derived large SiO_2 monoliths. *J. Sol–Gel Sci. Technol.* **6**, 203–17

Lee L-L, Tsai D-S 1994 Ion clustering and crystallization of sol–gel-derived erbium silicate glass. *J. Mater. Sci. Lett.* **13**, 615–7

Lochhead M J, Bray K L 1995 Rare-earth clustering and aluminum codoping in sol–gel silica: investigation using europium (III) fluorescence spectroscopy. *Chem. Mater.* **7**, 572–7

MacCraith B D, Ruddy V, Potter C, O'Kelly B, McGilp J F 1991 Optical waveguide sensor using evanescent wave excitation of fluorescent dye in sol–gel glass. *Electron. Lett.* **27**, 1247–8

Maeda H, Iwasaki M, Yasumori A, Yamane M 1990 Reduction of lead migration during drying of a gel. *J. Non-Cryst. Solids* **121**, 61–5

Matthews L R, Wang X -J, Knobbe E T 1994 Concentration effects on the luminescence behavior of europium (III) chloride and organoeuropium-doped silicate gel. *J. Non-Cryst. Solids* **178**, 44–51

Moreshead W V, Nogues J R, Krabill R H 1990 Preparation, processing and fluorescence characteristics of neodymium-doped silica glass prepared by the sol–gel process. *J. Non-Cryst. Solids* **121**, 267–72

Narang U, Bright F V, Prasad P N 1993 Characterization of rhodamine 6G-doped thin sol–gel films. *Appl. Spectrosc.* **47**, 229–34

Narang U, Gvishi R, Bright F V, Prasad P N 1996 Sol–gel-derived micron scale optical fibers for chemical sensing. *J. Sol–Gel Sci. Technol.* **6**, 113–9

Nogami M, Abe Y 1996 Properties of sol–gel-derived Al_2O_3–SiO_2 glasses using Eu^{3+} ion fluorescence spectra. *J. Non-Cryst. Solids* **197**, 73–8

Nogami M, Yamazaki T, Abe Y 1998 Fluorescence properties of Eu^{3+} and Eu^{2+} in Al_2O_3–SiO_2 glass. *J. Lumin.* **78**, 63–8

Nogues J-L R, Moreshead W V 1990 Porous gel–silica, a matrix for optically active components. *J. Non-Cryst. Solids* **121**, 136–42

Phalippou J, Wognier T, Zarzycki J 1984 Behavior of monolithic silica aerogels at temperatures above 1000 °C. In: Hench L L, Ulrich D R (eds.) *Ultrastructure Processing of Ceramics, Glasses and Composites*. Wiley, New York, Chap. 7, pp. 70–87

Pope E J A, Asami M, Mackenzie J D 1989 Transparent silica gel–PMMA composites. *J. Mater. Res.* **4**, 1018–26

Pope E J A, Mackenzie J D 1988 Nd-doped silica glass I: structural evolution in the gel state. *J. Non-Cryst. Solids* **106**, 236–41

Rabinovich E M, Wood D L, Johnson D W Jr, Fleming D A, Vincent S M, MacChesney J B 1986 Elimination of Cl_2 and H_2O in gel glasses. *J. Non-Cryst. Solids* **82**, 42–9

Reisfeld R, Eyal M, Gvishi R 1987 Spectroscopic behaviour of fluorescein and its di(mercury acetate) adduct in glasses. *Chem. Phys. Lett.* **138**, 377–83

Ruland G, Gvishi R, Prasad P N 1996 Multiphasic nanostructured composite: multi-dye tunable solid state laser. *J. Am. Chem. Soc.* **118**, 2985–91

Samuel J, Strinkovski A, Shalom S, Lieberman K, Ottolenghi M, Avnir D, Lewis A 1994 Miniaturization of organically doped sol–gel materials: a microns-size fluorescent pH sensor. *Mater. Lett.* **21**, 431–4

Satoh S, Susa K, Matsuyama Y, Suganuma T, Matsumura H 1997 Chlorination reduction in sol–gel-derived porous silica gels. *J. Non-Cryst. Solids* **217**, 22–9

Susa K, Matsuyama S I, Satoh S, Suganuma T 1982 New optical fiber fabrication method. *Electron. Lett.* **18**, 499–500

Tan W H, Shi Z-Y, Smith S, Birnbaum D, Kopelman R 1992 Submicrometer intracellular chemical optical fiber sensors. *Science* **258**, 778–81

Thomas I M, Payne S A, Wilke G D 1992 Optical properties and laser demonstration of Nd-doped sol–gel silica glasses. *J. Non-Cryst. Solids* **151**, 183–94

Wu F Q, Machewirth D, Snitzer E, Sigel G H Jr 1994 An efficient single mode Nd^{3+} fiber laser prepared by the sol–gel method. *J. Mater. Res.* **9**, 2703–5

Wu F Q, Puc G, Foy P, Snitzer E, Sigel G H Jr 1993 Low-loss rare earth doped single-mode fiber by sol–gel method. *Mater. Res. Bull.* **28**, 637–44

Zaitoun M A, Kim T, Lin C T 1998 Observation of electron–hole carrier emission in the Eu^{3+}-doped silica xerogel. *J. Phys. Chem.* **B102**, 1122–5

Zhao C F, Gvishi R, Narang U, Ruland G, Prasad P N 1996 Structures, spectra, and lasing properties of new (aminostyryl) pyridinium laser dyes. *J. Phys. Chem.* **100**, 4526–32

Zhao C F, He G S, Bhawalkar J D, Park C K, Prasad P N 1995 Newly synthesized dyes and their polymer/glass composite for one- and two-photon pumped solid-state cavity lasing. *Chem. Mater.* **7**, 1979–83

Zhou Y, Lam Y L, Wang S S, Liu H L, Kam C H, Chan Y C 1997 Fluorescence enhancement of Er-doped sol–gel glass by aluminum codoping. *Appl. Phys. Lett.* **71**, 587–9

A. Biswas, C. S. Friend and P. N. Prasad
State University of New York at Buffalo, New York, USA

Ceramic-modified High-temperature Polymers

Property enhancement through hybridization in the area of thermally stable polymers and ceramics is important for many applications in the electronics and aerospace industries. Preparation of such ceramers through the sol–gel process (Brinker 1988) allows systems to be designed that combine superior thermal stability, high refractive index, low thermal expansion coefficient, and a wide range of permittivity for ceramics, and to provide the toughness and crack deflection properties of polymers (Mark *et al.* 1995, Mark 1996). One of the challenges to obtaining significant improvements in these classes of polymers, however, is due to the fact that many of these polymers either possess superb thermal and mechanical properties to begin with or they are insoluble and infusible in organic solvents.

The unreactivity of such polymers that qualifies them as high-temperature, high-performance materials generally causes poor interfacial bonding between the polymers and the ceramic phases. As a consequence the molecular level combination of the two different components necessary to optimize the property combination may be difficult to achieve. Relatively few of the materials synthesized using these polymers have been exploited commercially. A brief account of the different strategies that have been used to combine thermally resistant polymers and ceramic phases using this process shall be discussed in this review.

The progress in heat-resistant polymers has been governed by the opposing requirements of thermal stability and processability. The structural requirement needed to attain high thermal stability means using the strongest chemical bonds in the polymer chain and resonance stabilization, and ensuring that these are no easy paths for rearrangement reactions. Polybonding is often utilized. Ring structures that have normal bond angles are preferred but to ease processability, thermally less stable flexible linking groups often have to be introduced between rings. Those having the least effect on stability are –COO–, –CONH–S–, $-SO_2-$, –O–, $-CF_2-$, and $-C(CF_3)_2-$.

Many thermally stable polymers, for example, polyesters, polyphosphazenes, aramids, and those containing heterocyclic rings: polyimides, polybenzoxazoles, polybenzobisthiazoles, and polybenzimidazoles have been used to combine with the inorganic components via the sol gel process. Silica is the ceramic phase most commonly generated in this sol–gel route to composites, but other ceramics such as titania (TiO_2), alumina (Al_2O_3), and zirconia (ZrO_2) have also been used. A combination of two or more oxides may also be used, as one type of oxide may improve one set of properties while the other improves different ones.

Morphology and phase separation control is critical in the generation of such hybrid materials. This problem has been solved by adding a coupling agent, functionalizing the oligomer or polymer chain ends, selecting polymers with appropriate backbone structures, or by using copolymers possessing functional groups that can interact with the growing inorganic oxide network. The relation between the structure of multiphase materials and their properties has been reviewed by many authors although there are severe limitations to theoretical models describing these parameters. In general the deformational behavior in such systems depends on the nature of the components, the volume fraction, size, shape, and orientation of the filler, and the state of adhesion between the filler and polymer matrix.

Radiation scattering, NMR, and other state-of-the-art techniques have been used to characterize these parameters. Generally these materials exhibit greater stiffness and hardness, although toughness is reduced. Highly transparent films, coatings and membranes have been produced using different means to control morphology. Preparation and properties of ceramers from some thermally resistant polymers are now described. Additional information is given in *Polymer–Ceramic Nanocomposites: Polymer Overview**.

1. Polyimides

The polyimides (PI) have become widely used as high performance polymers owing to their excellent thermal stability and toughness. However the thermal expansion coefficient (TCE), mechanical and dielectric properties, and so on, need to be strictly controlled for their application as circuit-printing films and semiconductor-coating materials in the microelectronics industry, and as adhesives in the aerospace industry. Several studies (Chujo 1996, Mascia 1995) have been carried out on the preparation of polyimide–metal oxide hybrid materials for this purpose through the sol–gel process. These materials have usually been produced by mixing solutions of pyromellitic dianhydride (PMDA), oxydianiline (ODA), and a metal oxide precursor, for example, tetraethoxysilane (TEOS), carrying out the hydrolysis and polycondensation of TEOS followed by an imidization process at high temperatures. Thin films were found to be transparent at low silica levels but were opaque at higher ones.

Despite the expected strong interactions through hydrogen bonds between the polyamic acid, the precursor for the PI, and hydroxyl groups from the silica phase, it was rather difficult to prevent phase separation. These systems have been compatiblized (Wang *et al.* 1994) with the selective use of thermally stable organically substituted alkoxysilanes, for example, aminophenyltrimethoxysilane, (aminoethylaminomethyl)-phenethyltrimethoxysilane, or 1-trimethoxysilyl-2-(*m*,*p*-chloromethyl)phenylethane, which undergo hydrolysis and polycondensation along with TEOS to form silica or polymeric silicates, while the groups at their other ends (such as the amino or chloro) provide secondary bonds with the PI chain.

Mascia and Kioul (1995) and Menoyo *et al.* (1998) have reported the use of glycidyl and isocyanate-type trialkoxysilanes to "couple" a polyamic acid to the silicate network prior to the condensation reactions for the formation of ceramer. By partial replacement of TEOS with dimethyldiethoxysilane and by varying the mixing conditions, morphologies ranging from interpenetrating networks to finely dispersed particles were obtained.

These results illustrate the role of the interconnected silica-rich particle within the polyimide-rich matrix in depressing the α-relaxations and reducing the TCE accordingly by an extent substantially greater than expected from the usual additivity rule.

Nandi *et al* (1991) used the concept of "site isolation" to reduce the silica particle size. Constraining phase separation between the polymer and inorganic clusters to a smaller scale provides optically transparent materials with improved properties. Masa-aki *et al.* (1992) introduced functional groups into the polymer backbone using ethoxysilylated diamines in place of ODA which provided bonding sites for the silica. The silica particle size was found to decrease and the modulus to increase as the number of bonding sites increased. Habib (2000) partially replaced ODA with 2,2-bis(3-amino-4-hydroxyphenyl)methane, a diamine monomer that has two pendant hydroxyl groups to develop extensive bonding between the PI and the silica network. Transparent hybrid films containing 70 wt.% of silica were obtained.

The relationships between various mechanical properties and the silica content are given in Fig. 1 The tensile modulus increased with an increase in silica content. Unlike the unbonded system where the tensile strength decreases with an increase in silica level these ceramers show slight increases and then retention of tensile strength up to much higher (25 .wt%) silica content. Extensive chemical bonding between the inorganic network and the polymer backbone solves the stress transfer problem thus reinforcing the hybrid films. The dynamical mechanical behavior of hybrid films also indicates (Fig. 2) the role of the bonded silica network in increasing the storage modulus and depressing substantially the α-relaxations. The SEM (scanning electron microscopy) micrograph of the unbonded ceramers (Fig. 3) shows sharp boundaries between the component phases, which become diffused in bonded ceramers (Fig. 4) showing interconnectivity between the phases. The thermal decomposition temperatures exhibited by these ceramers were at ~550 °C.

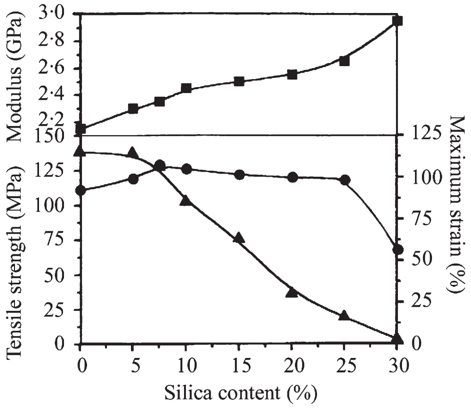


Figure 1
Comparison of the tensile properties as a function of the silica content in chemically bonded polyimide-silica ceramers. ● stress, ▲ strain.

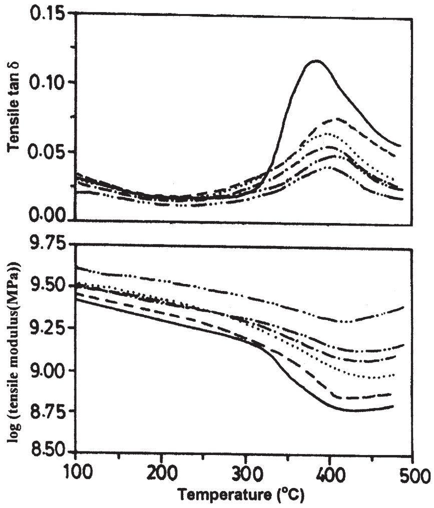


Figure 2
Temperature dependence of storage modulus and loss modulus for chemically bonded polyimide ceramers with silica wt%; ——, 0; ——, 10; · · · · ·, 15; –·–·, 20; –· ·–, 25; –· · ·–, 30.

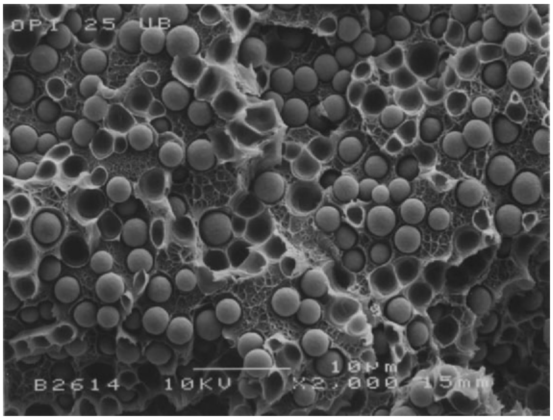


Figure 3
SEM micrograph showing discrete inorganic particles in nonbonded polyimide-silica ceramers.

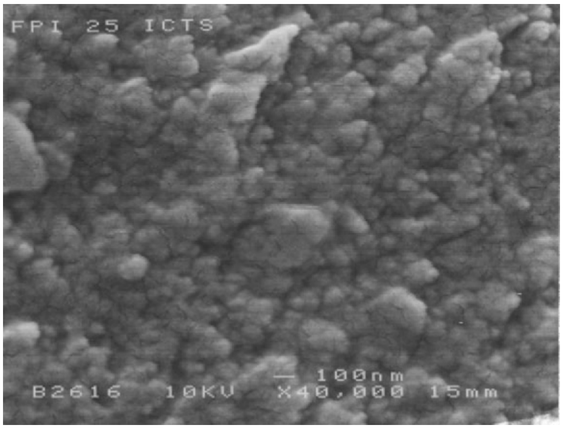


Figure 4
SEM micrograph showing diffused boundaries between the phases in chemically bonded polyimide-silica ceramers.

Ree *et al.* (1995) prepared PI nanocomposite films for use in fabrication of microelectronic devices by solution blending of polyamic acid and silica aerogels, and subsequently processing the PI film by conventional methods. The optical and dielectric properties were improved, whereas the interfacial stress and thermal expansion coefficient were significantly degraded by a large disturbance in the polymer chain in-plane orientation caused by silica aerogels despite their low thermal expansivity. This indicates that in the rigid type of polymer composites, the orientation of polymer chains still has an important effect on the physical properties. Nano-composite membranes based on PI and titania, and silica components with a higher permeability coefficient and significant improvement in preselectivity, have also been prepared using the sol–gel process. The dielectric constant was lowered substantially.

The sorption and permeation data for gases such as O_2, CO_2, H_2 and CH_4 have been correlated (Hu *et al.* 1997) to the morphology of the composite structure. Transparent PI films with large nonlinear optical (NLO) coefficient moieties have also been prepared (Jeng *et al.* 1992) where the linked alkoxysilane dyes and the polyimide form a semi-interpenetrating polymer network resulting in a substantial increase in the α-relaxation temperature and superb thermal stability of the poled order. The poled film of this PI-inorganic composite showed a large second-order NLO effect and excellent temporal stability, over 168 h at 120 °C.

2. *Polyaramids*

Ahmad and co-workers, (Ahmad and Mark 1998, Ahmad *et al.* 1997a, 1997b, 1998) have synthesized various types of aramid–silica composites. These aramids have been prepared from a terephthaloyl chloride and a mixture of *para*- and *meta*-phenylene diamines, with the former acting to stiffen the chains and the latter to provide some flexibility and tractability. In those hybrid materials where the aramid chains were not chemically bonded to the silica network, no significant increase in the mechanical strength of the polymer was observed. In another system, end-capping the aramid chains with aminophenyltrimethoxysilane developed chemical bonding between the aramid chains and the silica network. By end-linking the polymer chains with the silica network, transparency of the films improved greatly. The tensile strength was found to go through a maximum, whereas the elongation at break decreased and the tensile modulus increased with increasing silica content.

Interconnectivity of the two phases is thought to be the reason for improvements in mechanical properties. Increases in silica content beyond 10 wt.% in the matrix, however, caused a decrease in the tensile strength. Interfacial bonding through end-linking of the polymer chain can be achieved only up to a limit in linear polymers, because the higher the molecular weight of the polymer, the fewer the chain ends and the more limited the bonding capabilities of the chain. In some related materials, the average functionality of the monomer was increased in the synthesis simultaneously of linear–nonlinear polyaramid chains, which could be bonded into the inorganic network at more than just the two ends. Such chains have been prepared by the addition of benzenetricarbonyl chloride and terephthaloyl chloride (TPC) in a two-step process in which the TPC was first reacted with the diamines, with addition of the trifunctional monomer delayed until near the end of the reaction.

The linear parts of the chains give the desired strong intermolecular bonding, while the nonlinear chain ends provide more bonding points and thus better adhesion with the inorganic network. The introduction of silica raises the modulus, but the toughness and elongation to break are greatly reduced. The bonding between such polymers and *in situ* generated silica significantly reduce water absorption, and this improves the mechanical and dielectric properties of the matrix, particularly under humid conditions.

3. Polybenzoxazoles

The successful use of pendant functional groups in benzoxazole copolymers to provide chemical bonding between the polymer chain and the silica network has been made to produce thermally stable transparent ceramers. The copolymers were prepared by the polycondensation reaction of 4,4′-(2,2,2-trifluoro-1-(trifluoromethyl)ethylidene)bis(2-aminophenol) with a mixture of 4,4′-dicarboxybiphenyl ether and 5 hydroxyisophthalic acid or 5-phosphonoisophthalic acid. The tensile strength was found to depend on the nature of copolymer and the number of segments containing pendant functional groups in the chain; it increased with initial increase in the silica content and then decreased. The mechanical strength of the copolymer containing phospho acid groups was lower than those containing the hydroxyl groups (Chen *et al.* 1995).

4. Polyphosphazenes

The sol–gel synthesis of hybrid materials based on poly [bis ((methoxyethoxy) ethoxy) phosphazene] (MEEP) and silica has been reported. MEEP bears ethylene oxide oligomers as phosphorus side groups, which impart metal coordination ability to the polymer thus making it an ionic conductor. The low T_g of $-71°C$, has however limited the industrial development of MEEP. Chemical bonding of MEEP to a rigid inorganic network through the sol–gel process has been carried out (Guglielmi *et al.* 1996). Investigations on the thermal, mechanical, and electroconductive properties of the composite materials showed that these matrices presented improved mechanical and thermal features with respect to those of the original phosphazene. Abrasion-resistant transparent and antistatic MEEP–silica coatings have also been reported. Those prepared by the sol–gel process have been spin-coated on substrates and cured at ambient temperature or with moderate heating. The coating exhibits relatively steady conductivity with change in humidity, good transparency, toughness, hardness, and flexibility (Coltrain *et al.* 1990).

5. Polyesters

Transparent protective coatings from hydroxy-terminated polyesters and silica hybrids via the sol–gel process have been synthesized and evaluated for their performance on prefinish construction steel and aluminum. The interaction between the inorganic and organic domains occurs via hydroxyl end groups of polyesters and hydroxyl groups of silanols. These coatings combine flexibility, that is, upon deformation of the metal substrate after coating, with increased scratch resistance. The composition and molecular weight of the organic polymer, the ratio of the organic versus inorganic components, the cross-link density of the phases, solvents, catalysts, and curing conditions are some of the important parameters that control protein resistance of such coatings. Increases in Konig hardness and T_g was observed as the amount of inorganic component was increased (Frings *et al.* 1997).

6. Concluding Remarks

The inherent advantage of the sol–gel process and compatiblization of the organic and inorganic phases using organically substituted alkoxysilanes allows tailoring and improvement in materials properties. Many applications are still in the early stage of development, but transparent films, coatings, and membranes with high performance properties have been prepared and are being used in industry.

See also: High-temperature Stable Polymers$^{\#}$

Bibliography

Ahmad Z, Mark J E 1998 Biomimetic materials: recent development in organic-inorganic hybrids. *Mater. Sci. Eng.* **C-6**, 183–96

Ahmad Z, Sarwar M I, Mark J E 1997a Chemically bonded silica–polymer composites from linear and branched polyamides in a sol–gel process. *J. Mater. Chem.* **7** (2), 259–63

Ahmad Z, Sarwar M I, Wang S, Mark J E 1997b Preparation and properties of hybrid organic–inorganic composites prepared from poly(phenylene terephthalmide) and titania. *Polymer* **38** (17), 4523–9

Ahmad Z, Sarwar M I, Mark J E 1998 Thermal and mechanical properties of aramid-based titania hybrid composites. *J. Appl. Poly. Sci.* **70**, 297–302

Brinker C J 1988 *Better Ceramic Through Chemistry*. Materials Research Society, Pittsburgh, PA

Chen J P, Ahmad Z, Wang S, Mark J E, Arnold F E 1995 Preparation and mechanical properties of benzoxazole–silica hybrid materials. In: Mark J E, Lee C, Bianconi P A (eds.) *Hybrid Organic–Inorganic Composites*. American Chemical Society, Washington, DC

Chujo Y 1996 Organic–inorganic hybrid materials. *Curr. Opin. Solid State Mater. Sci.* **1** (6), 806–11

Coltrain B K, Ferrar W T, Landry C J 1990 Abrasion resistant phosphazene ether–metal oxide composites and their preparation. Patent WO 9011323

Frings S, van der Linde R, Ten H Meinema H 1997 Hybrid organic–inorganic coatings for prefinish construction steel and aluminium. In: Sarton L A J, Zeedijk H B (eds.) *Proc.*

Eur. Conf. Adv. Mater. Process Appl. Netherland Society for Material Science, Netherlands, Vol. 3, pp. 267–70

Guglielmi M, Brusatin G, Facchin G, Gleria M, De Jager R, Musiani M 1996 Hybrid materials based on metal oxides and poly(organophoshazenes). *J. Inorg. Organomet. Polym.* **6** (3), 221–36

Habib R 2000 Study on the structure/property relationships in polyimide–silica hybrids with extensive interphase bonding prepared through sol-gel process. PhD Thesis, Saarbrucken University, Germany

Hu Q, Marand E, Dhingra S, Fritsch D, Wen J, Wilkes G 1997 Polyamid-imide TiO_2 nano-composite gas seperation membranes: fabrication and characterization. *J. Membr. Sci.* **135** (1), 65–79

Jeng R J, Chen Y M, Jain A K, Kumar J, Tripathy S K 1992 Stable second order nonlinear optical polyimide-inorganic composite. *Chem. Mater.* **4**, 1141–4

Joly C, Goizet S, Schrotter J C, Sanchez J, Escoubes M 1997 Sol–gel polyimide–silica composite membrane: gas transport properties. *J. Membr. Sci.* **130** (1–2), 63–74

Mark J E 1996 The sol–gel route to inorganic–organic composites. *Heterogeneous Chem. Rev.* **3**, 307–26

Mark J E, Lee C Y-C, Bianconi P A (eds.) 1995 *Hybrid Organic–Inorganic Composites*. American Chemical Society, Washington, DC

Masa-aki K, Atsushi M, Yoshitake L, Yoshio I, Morikawa A, Iyoku I, Kakimoto M, Imai Y 1992 Preparation of new polyimide–silica hybrid materials via the sol–gel process. *J. Mater. Chem.* **2** (7), 679–90

Mascia L 1995 Development in organic–inorganic polymeric hybrids: ceramers. *Trends Polym. Sci.* **3** (2), 61–6

Mascia L, Kioul A 1995 Influence of siloxane composition and morphology on properties of polyimide-silica hybrids. *Polymer* **36** (19), 3649–59

Menoyo J D C, Mascia L, Shaw S J 1998 Compatibilization mechanism of polyimide–silica hybrids with organofunctional trialkoxysilanes. *Mater. Res. Soc. Symp. Proc.* Molecular-level ceramic polymer composites: synthesis of polymer-trapped silica and titania nanoclusters. **520**, 239–67

Nandi M, Conklin J A, Salvati H Jr, Sen A 1991 Molecular-level ceramic polymer composites: synthesis of polymer-trapped silica and titania nanoclusters. *Chem. Mater.* **3**, 201–6

Ree M, Goh W H, Kim Y 1995 Thin films of organic polymer composites with inorganic aerogels as dielectric materials. Polymer chain orientation and properties. *Polym. Bull.* **35**, 215–22

Wang S, Ahmad Z, Mark J E 1994 Polyimide–silica hybrid materials having interfacial bonding through use of a sol–gel technique. *Macromol. Rep. A* **31**, 411–9

Z. Ahmad
Quaidi-i-Azam University, Islamabad, USA

Ceramics and Glasses, Sol–Gel Synthesis of

The sol–gel process represents an elegant chemical route that can be used for the low-temperature synthesis of single- or multiple-component ceramic materials (including glasses) in the form of thin solid films, ultrafine powders, high surface area porous materials, dense abrasive minerals, and continuous ceramic and glass fibers. The sol–gel process involves the preparation of a precursor mix (a sol or a solution), which is ultimately converted into a final product via a stage that may involve drying, chemical reactions, gelation of the precursors, and curing (often using heat). The major advantages associated with the sol–gel processes include lower processing temperatures, high levels of purity, control of dopant concentrations, and the ability to synthesize multicomponent compositions in different product forms.

For some applications the cost of the starting materials used in sol–gel processing may be relatively high, yet it is used commercially for the preparation of continuous ceramic fibers, abrasive minerals, optical fibers, ceramic powders, and thin films. This article first describes and clarifies the different terms used in the area of sol–gel processing. This is followed by an overview of the sol–gel process that includes a description of the precursors used and the chemical reactions involved. In the final section an attempt is made to capture and highlight some of the applications of the sol–gel process.

1. Terminology

The term *sol* refers to a dispersion of colloidal particles of one phase in a fluid medium. For example, when colloidal particles are suspended in air, water, or an organic liquid, the terms aerosol, hydrosol, and organosol, respectively, are used. The particles in a sol, by definition, are colloidal, i.e., at room temperature they remain dispersed by virtue of Brownian motion. This usually requires the average particle size to be less than 1000 nm (1 μm). In practice, in the sols used in sol–gel processing, the average particle size of precursors used can consist of somewhat larger particles (a few micrometers in diameter) and, as is explained below, *solutions* can also be used. An important and fundamental distinction between a *sol* and a *solution* is that a true solution is thermodynamically stable. A sol, however, is a dispersion and is therefore thermodynamically unstable. The term *gel* is defined here as a material consisting of a three-dimensional network of a solid phase interwoven with an entrapped and immobilized continuous liquid phase. The sol–gel process has also sometimes been referred to as the *solution sol–gel* (SSG) process. This term is more appropriate if the precursors comprise a true solution that undergoes chemical reactions that lead to formation of a sol of colloidal species, which in turn transform into a gel.

The terms *colloidal sol–gel* and *polymeric* (*alkoxide*) *sol–gel* are used to describe sol–gel processes in which dispersions of ultrafine particles of ceramics or solutions of metal alkoxides (or other chemicals) are used, respectively. The terms *sol precipitation* or *chemical* *precipitation* have been used to describe chemical processes in which ultrafine particles of a ceramic material are formed by chemical reactions occurring in a solution of precursors. Ceramic powders can also be formed from curing of a gel obtained from colloidal or polymeric alkoxide precursors. The term *metalorganic decomposition* (MOD) has also been sometimes used interchangeably with the term *sol–gel*. MOD is actually a solution decomposition process and is similar to the spray pyrolysis process. Typically in MOD, long hydrocarbon chain metalorganic precursors are dissolved in organic solvents and the resulting solutions are then thermally decomposed to form thin films or powders. This article describes colloidal and polymeric sol–gel processes used for the synthesis of thin films, continuous fibers, powders, and large vitreous silica monoliths.

2. *Overview of the Sol–Gel Process*

In sol–gel processing of ceramics and glasses, chemical precursors such as solid colloidal particles and/or solutions of inorganic or organic compounds of metals are used to form a precursor mix, which is usually a sol of ceramic particles or a solution containing other dissolved chemicals (Fig. 1(a)). Colloidal precursors are often preferred for larger volume applications since they allow much higher levels (up to ~50 wt.%) of solids loading. Polymeric alkoxide precursors are more suited for thin-film deposition applications. The precursor mix is then formed into the desired shape such as thin films (Fig. 1(b)), ultrafine particles/powders, a bulk gel (Fig. 1(c)), or continuous fibers (Fig. 1(d)).

During this conversion the precursor mix transforms into a more solid-like material as a result of evaporation of the solvent, dehydration, and chemical cross-linking between the solid particles or dissolved precursors. The sol-derived "wet" thin film, gel, powder, or fiber is then cured to yield a final product. The curing process may involve simple drying (including drying under supercritical conditions), calcination, and sintering steps. During the curing step, the precursors may react with each other to form chemical bonds that are characteristic of a ceramic or a glass. For example, a colloidal silica sol or a silicon alkoxide dissolved in an alcohol may be used

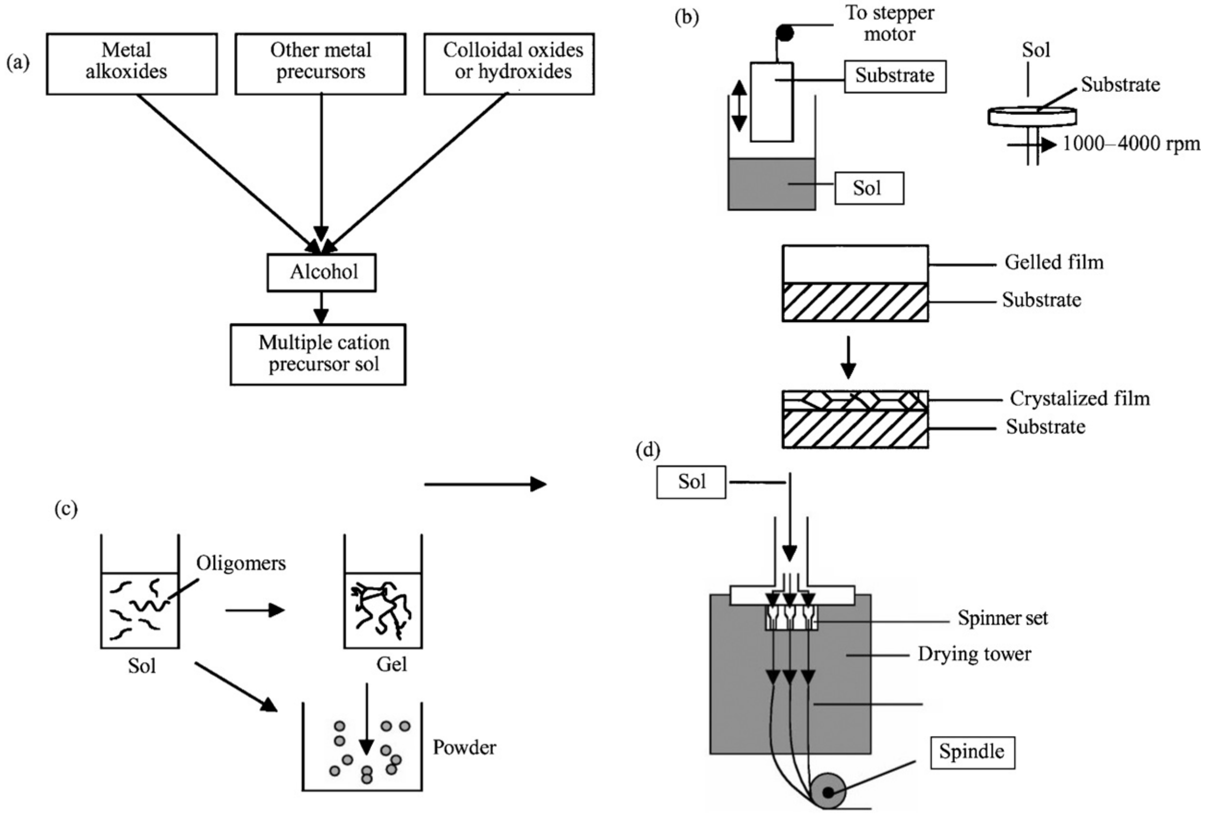

(a)
Metal alkoxides
Other metal precursors
Colloidal oxides or hydroxides
Alcohol
Multiple cation precursor sol
(b)
To stepper motor
Substrate
Sol
Sol
Substrate
1000–4000 rpm
Gelled film
Substrate
Crystalized film
Substrate
(c)
Oligomers
Sol
Gel
Powder
(d)
Sol
Spinner set
Drying tower
Spindle

Figure 1
Overview of the sol–gel process and some of applications for the synthesis of bulk gels, powders, thin films, and fibers.

to create an essentially inorganic silica (SiO_2) thin film or a silica fiber. The transformation of a sol or solution into an essentially inorganic crystalline ceramic or glass, via a gel-like intermediate, can be often accomplished at temperatures that are significantly lower that those necessary in "traditional" processing. This is because the precursors used are intermixed at a molecular level and can react through chemical pathways to form a network of chemical bonds that are characteristic of completely inorganic materials. Therefore, in sol–gel processing high processing temperatures, used typically to induce long-range diffusion of atoms, are often not needed. Temperatures used in the curing step may range from $\sim 300\,^{\circ}C$ to $1500\,^{\circ}C$, depending on the composition of the material being prepared. The concept of using *seeds* or *templates* (Hong and Messing 1999), nucleating agents as well as heterometallic precursors, has also been shown to be useful in lowering the temperatures needed for formation and densification of the desired phases. This lowering of the synthesis and processing temperatures is one of the most important advantages of using the sol–gel process for the synthesis of ceramics and glasses.

The chemical nature of the sol–gel process provides additional advantages. Since the degree of mixing of two or more phases is dictated by the nature of the chemical precursors (molecular vs. colloidal), the sol–gel process, by proper design, can be used to prepare thermodynamically stable as well as metastable phases of materials. This chemical control of the degree of homogeneity of the precursors also allows the controlled introduction of selected dopants into the sol–gel ceramic material. Finally, another key advantage of the sol–gel process is that since mostly synthetic precursors are used rather than naturally occurring minerals, the purity of sol–gel-derived fibers, powders, and thin films can be exceptionally high.

3. Precursors for the Sol–Gel Process

In sol–gel processing, metal alkoxides ($M(OR)_n$; M^{n+} = metal cation of valence n and OR = alkoxy group) or other metalorganic chemicals and/or colloidal metal oxides or hydroxides can be used as precursors. The metalorganic precursors are characterized by the presence of M–O bonds and are different from organometallic precursors, which are characterized by the presence of M–C bonds. For example, for the preparation of silica (SiO_2) films, an alkoxide of silicon, e.g., tetraethylorthosilicate ($Si(OC_2H_5)_4$), dissolved in an alcohol or a colloidal silica sol can be used as the starting material. Alkoxides of many individual metals are available commercially. Many metal alkoxides, especially those that are liquids, can be obtained in exceptionally high levels of purity (>99.99%, metals, basis). This purity, combined with the volatile nature of certain metal alkoxides, enables these precursors to be used not only in the sol–gel process but also in processes such as chemical vapor deposition (CVD). In certain cases it is also possible to use heterometallic alkoxides. Such heteroalkoxides have been used to advantage in the synthesis of binary oxides such as mullite ($3Al_2O_3 \cdot 2SiO_2$) and spinel ($MgO \cdot Al_2O_3$).

Precursors such as acetylacetonates and acetates or metal alkoxides molecularly modified with these functionalities have also been used for the sol–gel synthesis of ceramics. These precursors are useful for metals whose alkoxides are either insoluble (e.g., some of the group 2 elements) or extremely reactive (e.g., group 1 elements). Metal nitrates have also been used but are not preferred since they tend to crystallize during the subsequent drying steps. Precursors such as metal chlorides may be used but contamination owing to remnant anions may not be acceptable. Partially neutralized, soluble salts known as "basic" metal salts are also employed as sol–gel precursors. Basic metal salt solutions generally are comprised of complex mixtures of polynuclear metal cations (as in the case of aluminum, zirconium, titanium, iron, chromium, and beryllium) and concentrated basic metal salt mixtures have been used to produce ceramic metal oxide fibers. Sols of colloidal particles of silica, alumina, titania, and other materials are also available commercially.

4. Chemistry of the Sol–Gel Process

The chemical basis for the *polymeric sol–gel* process is similar to the well-known polymerization process. Alkoxide molecules exist as low molecular weight species, usually monomers or small oligomers, depending on the particular alkoxides used, the solvent, and other factors. The polymerization of these species occurs via hydrolysis of the M–OR moiety and polycondensation reactions involving the resulting M–OH group. Water acts as the "initiator" and is usually externally added. Water can also be generated *in situ* by condensation reactions such as the formation of esters by the condensation of carboxylic acids and alcohols. Acids or bases are used as catalysts for these reactions. The initial viscosity of the sol derived from polymeric precursors is rather low and in many cases is similar to that of water ($\sim 1\,mPas$). If the polycondensation reactions are allowed to occur to a significant extent, the original sol can transform into a gel.

The reaction of metal alkoxides with water is known as the hydrolysis reaction:

$$M(OR)_x + xH_2O \rightarrow M(OH)_x + xROH \qquad (1)$$

Note that Eqn. (1) represents a complete hydroxy–alkoxy exchange, i.e., all the alkoxy groups have been shown to be replaced by hydroxy groups. This may or may not occur depending on the relative concentration of water and other conditions. The partial hydrolysis can be represented by

$$\mathrm{M(OR)}_x + y\mathrm{H_2O} \rightarrow \mathrm{M(OH)}_y\mathrm{(OR)}_{x-y} + y\mathrm{ROH} \quad (2)$$

One of the implications of the hydrolysis reaction is that proper care must be taken in handling, packaging, and using metal alkoxides. Even traces of moisture are often sufficient to cause prehydrolysis and undesirable precipitation of hydroxides.

The next step is the following condensation reactions:

$$\mathrm{M{-}OR + M{-}OH \rightarrow M{-}O{-}M + R{-}OH} \quad (3)$$

$$\mathrm{M{-}OH + M{-}OH \rightarrow M{-}O{-}M + H{-}OH} \quad (4)$$

In Eqns. (3) and (4) the remaining functionalities on the metal cations are not shown. Notice that Eqn. (4) also can be used to describe network formation between hydroxyl-functional, colloidal metal oxide precursors. For example, colloidal silica particles can undergo gelation through reactions such as that shown in Eqn. (4). The products of condensation reactions may be alcohol, water, or both, depending on the alkoxide used. Thus metal–oxygen–metal (M–O–M) bonds can form during the condensation reactions that occur using either alkoxide or colloidal oxide precursors. These reactions ultimately lead to the formation of a network consisting of metal–oxygen bonds. An important feature of these reactions is that they can occur at low temperatures ($\sim < 100\,^{\circ}\mathrm{C}$). This is one of the reasons that metal alkoxides and colloidal oxide particulates represent the precursors of choice for gel processing of oxide ceramics and glasses. Cross-condensation reactions, i.e., reactions between species from two or more cations, are also possible:

$$\mathrm{M^1(OR) + M^2(OH) \rightarrow M^1{-}O{-}M^2 + R{-}OH} \quad (5)$$

$$\mathrm{M^1(OH) + M^2(OH) \rightarrow M^1{-}O{-}M^2 + H{-}OH} \quad (6)$$

Again, note that Eqn. (6) can also represent linking of colloidal particles of different precursors, e.g., colloidal silica and colloidal alumina. In principle, a solution containing different alkoxides could be used to construct a network of many cations. In practice, however, differential hydrolysis and polycondensation rates of different alkoxides pose a problem. Also, owing to the insolubility or other factors (such as extreme reactivity, cost, etc.) certain metals may be introduced in the form of other precursors that do not exhibit hydrolysis and polycondensation reactions. Different approaches including prehydrolysis of the more slowly reacting metal alkoxides as well as stabilization of highly reactive metal alkoxides by the molecular modification process have been developed to tackle this important problem (Phulé and Khairulla 1990).

5. *Applications of the Sol–Gel Process*

The historical routes and applications of the sol–gel process for synthesis of ceramic materials and glasses have been traced to the archives dating back to the sixteenth century and perhaps even before (Wood and Dislich 1995). Although the term "sol–gel" was to be coined later, technology identical to that now known as sol–gel was used to produce adsorbents, catalysts, and ceramic coatings in the period 1920–1950. Roy (1956) reported on the use of this process for a controlled chemical synthesis of multicomponent minerals. In the early 1960s the sol–gel process found application in the preparation of thorium oxide and uranium dioxide microspheres used in the nuclear industry and it was during this period that sol–gel was given its name. Since the early years of the sol–gel process, considerable technological development has occurred in this area. Noteworthy examples of commercial applications of sol–gel technology include, but are not limited to, the areas of optical fibers, refractory and reinforcing ceramic fibers, abrasive minerals, and ceramic and ceramic-like coatings.

The first practicable sol–gel process for making high-quality silica optical fibers was reported by researchers at Lucent Technology (MacChesney *et al.* 1999). More specifically, in this application the sol–gel process is used for the production of large overcladding silica tubes. The material from these tubes ultimately results in about 90% of the mass of the fiber and this application requires exceptionally pure silica glass. The sol–gel process developed for optical fiber production is different from the schematic shown in Fig. 1(d). Researchers have used colloidal silica (Degussa OX-50, surface area $50\,\mathrm{m^2\,g^{-1}}$) that is dispersed in an aqueous solution of tetramethylammonium hydroxide, a polymer, and a lubricant. The stabilized sol of silica particles is centrifuged to remove impurity particles and de-aired in vacuum.

An ester such as methyl formate is added to the sol (to lower the pH) and then the sol is pumped into a mold to form the so-called overcladding tubes. Gelation occurs within 10 minutes, following which the gel body is removed and dried in a controlled atmosphere. Dry tubes are heated first to remove residual water and volatile organics and then heating in a chlorine-containing atmosphere is also used to remove hydroxyl ions, transition metal impurities, and refractory particles. This is followed by sintering at 1500–1550 °C to produce a vitreous cylinder. Optical fiber preforms are then fabricated by collapsing the tube over a vapor-deposited core rod. High-quality optical fibers are then drawn from such optical fiber performs. The mass of such overcladding tubes that can be prepared using the sol–gel process approaches 20 kg.

Other examples of commercial applications of the sol–gel process include the development of the Nextel fibers and Cubitron abrasive grains. Nextel fibers are prepared by extruding, drying, and firing a concentrated sol containing a basic aluminum salt mixture along with additives such as colloidal silica, organic polymers, and, in some cases, a boron oxide precursor. In general, the Nextel fibers are polycrystalline, transparent, and can be woven or braided for different engineering applications. Nextel 312 fibers based on a ceramic consisting of 62% Al_2O_3, 14% B_2O_3, and 24% SiO_2, are processed using a colloidal sol–gel process. Nextel 610 fibers are high-modulus, α-alumina ceramic fibers having an exceedingly fine grain size (average grain size less than 2000 nm). In the production of these fibers nucleating agents are used to reduce the average grain size. A micrograph of α-alumina fine-grained fibers produced with the use of nucleating agents is shown in Fig. 2(a). For comparison, a micrograph of the relatively coarse-grained and porous microstructure of fibers prepared without using nucleating agents is shown in Fig. 2(b). Such sol–gel-derived ceramic fibers and products made using these (Fig. 3) find applications in ceramic fiber-reinforced metal matrix composites and in reinforcing ceramics and plastics.

While sol–gel processes have been used for many years to produce porous ceramic materials that have been used as adsorbents and catalyst supports, in the 1970s researchers at 3M began to develop technologies to form dense and tough ceramic materials via the sol–gel approach. This effort resulted in the development of a number of sol–gel-derived abrasive minerals, including materials based on α-alumina–spinel mixtures, superfine-grained aluminas, and alumina–rare earth aluminates. The synthesis of these materials involves first forming a concentrated sol of nanocrystals of aluminum oxyhydroxide, boehmite.

The boehmite that is used is a synthetic product that is formed via hydrolysis of an aluminum alkoxide generated by the oxidation of aluminum metal by alcohol at elevated pressure and temperature. The concentrated boehmite sol is gelled either by raising the ionic strength of the sol or by heating. Modifying agents such as phase-forming precursors or nucleating agents for enhancing the formation of α-alumina can be added to the sol prior to gelation. To form the ceramic abrasive grain, the boehmite gel is dried, crushed to size, calcined, and fired to full density. High-performance abrasive grains of this general type are manufactured and distributed worldwide on a very large scale by 3M Company.

In many familiar applications, sol–gel-derived products play an important role that is often not readily recognized. For example, precipitated silicas are used as a filler material in the manufacture of automobile tires. Other products such as paints and toothpastes also often make use of colloidal silica particles. Some of the other applications of sol–gel-derived products include antireflective (AR) coatings, spin-on glasses for dielectric planarization, ferroelectric thin films, electrochromic coatings, and thin layers of transparent oxide conductors. Also of interest is the area of sol–gel-like processes for polysiloxane–silica coatings on clear plastics for improving abrasion resistance of ophthalmic products such as eyeglasses and contact lenses.

Much effort in the area of sol–gel processing has been devoted to the synthesis and processing of oxide

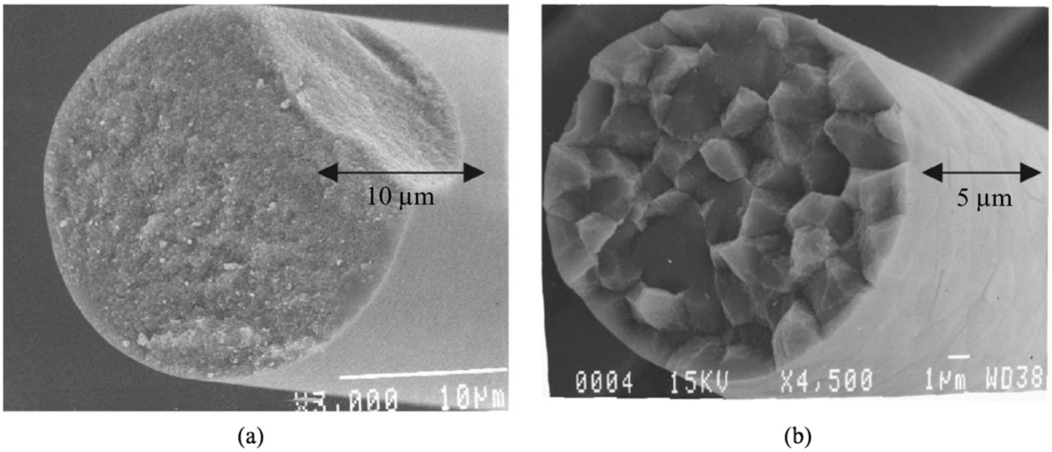


Figure 2
Microstructure of sol–gel-derived α-alumina ceramic fibers: (a) superfine-microstructure, dense alumina fiber after firing with nucleating agent; (b) coarse-microstructure, porous alumina fiber after firing with no nucleating agent.

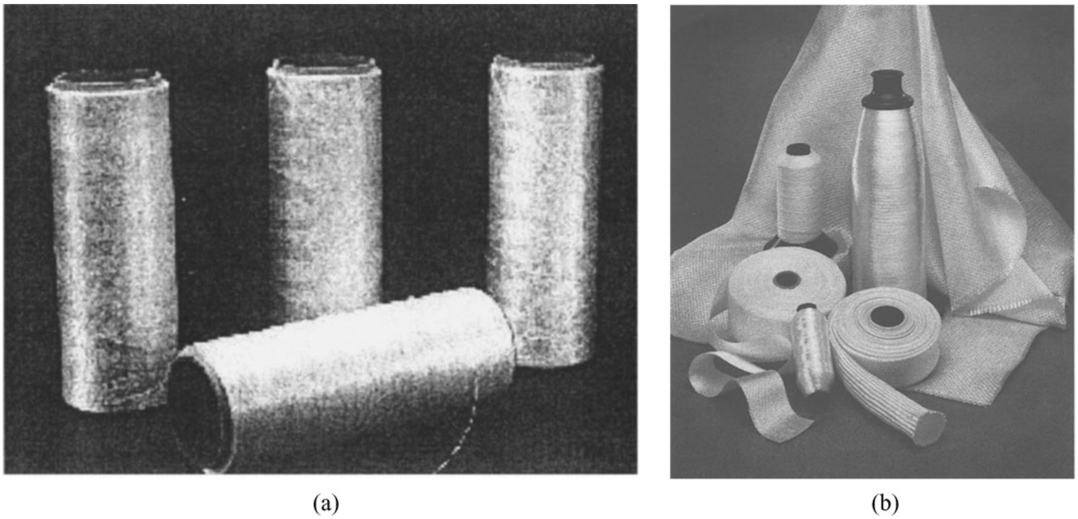


Figure 3
Examples of commercial products prepared using sol–gel processing: (a) yarn of Nextel 312 fibers; (b) insulating products manufactured using these fibers.

ceramics and glasses. However, the gel synthesis of nonoxide ceramics (nitrides, sulfides, carbides, etc.) has also received considerable attention. For more detailed information on the sol–gel process the reader is referred to the bibliography below. Internet sites such as at http//www.solgel.com may also yield useful and updated information.

Bibliography

Bradley D C, Mehrotra R C, Gaur D P 1978 *Metal Alkoxides*. Academic Press, New York

Brinker C J, Scherer G W 1990 *Sol–Gel Science: The Physics and Chemistry of Sol–Gel Processing*. Academic Press, New York

Erickson D D, Wood T E, Wood W P 1999 Historical development of abrasive grain. *Am. Ceram. Soc., Ceram. Trans.* **95**, 73–86

Hench L L 1998 *Sol–Gel Silica—Properties, Processing, and Technology Transfer*. Noyes, Park Ridge, NJ

Hong S H, Messing G L 1999 Development of textured mullite by templated grain growth. *J. Am. Ceram. Soc.* **82** (4), 867–72

Klein L C 1988 *Sol–Gel Technology for Thin Films, Fibers, Preforms, Electronics, and Specialty Shapes*. Noyes, Park Ridge, NJ

MacChesney J B, Johnson Jr D W, Bhandarkar S, Bohrer M, Fleming J W, Monberg E M, Trevor D J 1999 Sol–Gel process for optical fiber manufacture. *Am. Ceram. Soc., Ceram. Trans.* **95**, 65–72

Phulé P P, Khairulla F 1990 Molecularly modified alkoxide precursors for chemical synthesis of dielectric ceramics. In: Zelinski B J, Brinker C J, Clark D E, Ulrich D R (eds.) *Better Ceramics Through Chemistry IV*. Materials Research Society, Vol. 180, pp. 527–32

Pierre A C 1998 *Introduction to Sol–Gel Processing*.

Roy R 1956 Aids in hydrothermal experimentation: II. Methods of making mixtures of both "dry" and "wet" phase equilibrium studies. *J. Am. Ceram. Soc.* **39** (4), 145–6

Wood T E, Dislich H 1995 An abbreviated history of sol–gel technology. *Am. Ceram. Soc., Ceram. Trans.* **55**, 3–23

P. P. Phulé[a] and T. E. Wood[b]
[a] *University of Pittsburgh, Pennsylvania, USA*
[b] *3M Corporation, St. Paul, Minnesota, USA*

Ceramics: Whisker Toughening

Whiskers of ceramic materials, with their potential high strength coupled with high stiffness, refractoriness, and low theoretical density, originally were considered as agents for reinforcing metallic alloys in the early 1960s. In the early 1980s, significant breakthroughs were made in reinforcing ceramics with ceramic-based whiskers. The first successful development involved an aluminum oxide matrix reinforced with microscopic single-crystal silicon carbide whiskers (Becher and Wei 1984). The whiskers used are typically less than 1 μm in diameter—much smaller than a human hair ($\geqslant 30\,\mu\mathrm{m}$ in diameter). The whiskers serve to toughen the ceramic, acting much like rebar does in concrete or nails used to hold two pieces of wood together. To do this the whiskers must remain intact as a crack advances through the composite and they must bridge across the crack, serving to hold the crack together until they eventually fracture

Table 1
Mechanical properties of fine-grained (2 μm) alumina composites.

	Flexure strength (MPa)			Fracture toughness (MPa m$^{1/2}$)		
Temperature (°C)	22	1000	1200	22	1000	1200
With no reinforcement	$\leqslant$550			$\leqslant$3		
With 10 vol.% SiC whiskers	600			6		
With 20 vol.% SiC whiskers	800	775	550	8.4	8.1	10

or are pulled out of the matrix. For this to occur, the interface between a whisker and the ceramic matrix must separate (debond) before the whisker can fracture, as the goal is for the crack front to go around the whiskers rather than through them. A number of factors influence the properties of whisker-reinforced ceramics including the size (length, diameter, and aspect ratio) and morphology (e.g., straightness, surface smoothness) of the whiskers, and matrix characteristics and processing parameters (Becher *et al.* 1988, Becher 1991, Homeny *et al.* 1986).

Whisker-reinforced alumina (WRA) composites quickly found application in high-speed machining of metallic alloys (Smith 1986) where their wear and fracture resistance allowed for higher precision machining and lower requirements for lubrication (reducing waste by lowering requirements for solvent cleaning). WRA cutting tools are used in high-speed machining operations involving hard materials such as nickel-based superalloys used in turbine engines. Other applications include machining of fiberglass epoxy composites (e.g., instrument boards), and in woodworking and construction (e.g., longer service life with resistance to wear and corrosive attack associated with modern particle and chipboard materials). Attempts to use whisker reinforcement of other ceramics have not been as successful; however, whisker-reinforced silicon nitride composites (Campbell *et al.* 1990, Sajgalik and Dusza 1989) have been developed and are used in similar machining applications. The known annual worldwide sales of WRA components are reported to have a total value in excess of US$50 million. Some of the properties that contribute to the application of WRA composites are summarized below. The reader is referred to the bibliography for further information.

1. *Properties of SiC WRA Composites*

With careful selection of the SiC whiskers and the processing conditions, WRA composites, in comparison to nonreinforced alumina ceramics, can exhibit substantial increases in fracture toughness and fracture strength (Becher and Wei 1984, Becher *et al.* 1988, Becher 1991), which are also retained as the test temperature is increased (Becher and Tiegs 1988). Furthermore, the improvements are also found in thermal shock resistance (Tiegs and Becher 1987, Becher and Warwick 1993), crack growth resistance when exposed to both static and cyclic loading (Becher *et al.* 1990, Dauskardt *et al.* 1993, Han and Suresh 1989), and creep resistance (Lin and Becher 1990, 1991) in SiC WRAs. Examples of some of the improvements in fracture strengths and toughnesses that have been achieved are listed in Table 1.

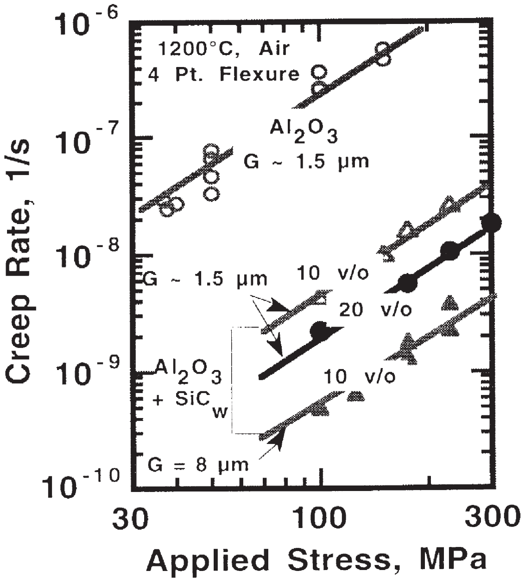


Figure 1
Resistance of alumina to creep at elevated temperatures increases significantly with the addition (e.g., 10 and 20 vol.%) of SiC whiskers. Increasing the grain size (*G*) of the alumina matrix (see Lin and Becher 1990, 1991) can enhance this further.

An added attractive feature is the greater creep resistance of WRA composites, as shown in Fig. 1. The creep resistance in air is limited by a weakening effect caused by oxidation of the SiC whiskers to form SiO_2, which reacts with alumina to form a glass. Because this is a thermally activated process, WRA composites are typically quite oxidation resistant until temperatures are in excess of 1100 °C and this oxidation-driven reaction can also be limited by

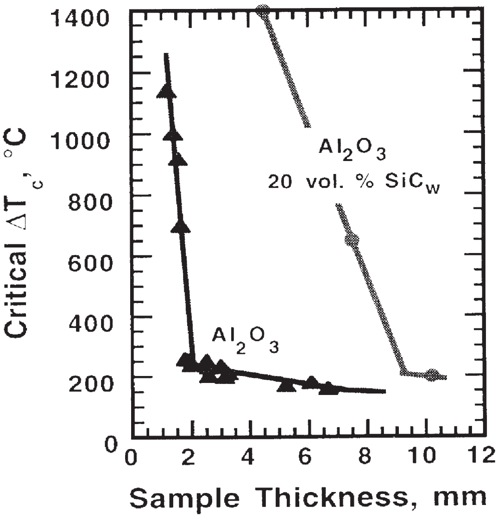


Figure 2
Thermal shock resistance is substantially increased by whisker reinforcement. Even sections that are several millimeters thick do not undergo strength losses after quenching from several hundred degrees centigrade into a boiling water bath (Becher and Warwick 1993).

lowering the SiC content. The combination of crack growth (Becher *et al.* 1990, Dauskardt *et al.* 1993, Han and Suresh 1989) and creep resistance contribute to the superior performance of WRA composites in high-speed machining of superalloys where high temperatures in the alloy at the tool tip are needed to enhance the cutting speed and surface finish.

Ceramics that have low thermal conductivity are susceptible to failure when rapidly cooled or heated (i.e., thermally shocked) as a result of the low toughness and strength and high Young's modulus. A convenient measure of the degree of thermal shock resistance is reflected in the temperature from which a sample is quenched into a liquid held at a fixed temperature that results in a recognizable loss in strength. This is known as the critical (quenching) temperature (ΔT_c). As a result of heat transfer effects during the quenching process, ceramics are more readily damaged as the smallest dimension of the sample increases. Alumina ceramics are no exception; however, the addition of SiC whiskers, which increases both the toughness, strength, and thermal conductivity, is found also to result in a far greater thermal shock resistance as the sample size increases (Fig. 2).

Acknowledgment

The research was sponsored by the US Department of Energy, Office of Basic Energy Sciences, Division of Materials Sciences and Engineering, under contract DE-AC05-00OR22725 with UT-Battelle, LLC.

Bibliography

Becher P F 1991 *J. Am. Ceram. Soc.* **74** (2), 255–69
Becher P F, Angelini P, Warwick W H, Tiegs T N 1990 *J. Am. Ceram. Soc.* **73** (1), 91–6
Becher P F, Hsueh C H, Angelini P, Tiegs T N 1988 *J. Am. Ceram. Soc.* **71** (12), 1050–61
Becher P F, Tiegs T N 1988 *Adv. Ceram. Mater.* **3** (2), 148–54
Becher P F, Warwick W H 1993. In: Schneider G A, Petzow G (eds.) *Thermal Shock and Thermal Fatigue Behavior of Advanced Ceramics*. Kluwer, Dordrecht, The Netherlands, pp. 37–48
Becher P F, Wei G C 1984 *J. Am. Ceram. Soc.* **67** (12), C267–9
Campbell G H, Rühle M, Dalgleish B J, Evans A G 1990 *J. Am. Ceram. Soc.* **73** (3), 521–30
Dauskardt R H, Yao D, Dalgleish B J, Becher P F, Ritchie R O 1993 *J. Mater. Sci.* **28**, 3258–66
Han L X, Suresh S 1989 *J. Am. Ceram. Soc.* **72** (7), 1233–8
Homeny J, Vaughn W L, Ferber M K 1986 *Am. Ceram. Soc. Bull.* **65** (2), 333–8
Lin H T, Becher P F 1990 *J. Am. Ceram. Soc.* **73** (5), 1378–81
Lin H T, Becher P F 1991 *J. Am. Ceram. Soc.* **74** (8), 1886–93
McCluskey P H, Williams R K, Graves R S, Tiegs T N 1989 *Ceramic Technology for Advanced Heat Engines Project Semiannual Progress Report*, ORNL/TM-11116. Oak Ridge National Laboratory, Oak Ridge, TN, pp. 274–83
Sajgalik P, Dusza J 1989 *J. Eur. Ceram. Soc.* **5**, 321–6
Smith K H 1986 *Machine and Tool Blue Book* **81** (1), 71–2
Taylor R E, Groot H 1983 *Thermophysical Properties of SiC-Reinforced Al$_2$O$_3$*, PRL 692. Properties Research Laboratory, West Lafayette, IN
Tiegs T N, Becher P F 1987 *J. Am. Ceram. Soc.* **70** (5), C-109–11

P. F. Becher
Oak Ridge National Laboratory, Tennessee, USA

Cermets and Hardmetals

The word "cermet" is a term that designates a material made of the combination of a metallic alloy and a ceramic. A wide range of materials is covered by this definition. In agreement with common use, it is more appropriate to define as cermets those ceramic–metal composites that contain a predominant ceramic phase (carbides, nitrides, borides, silicides, oxides, and eventually ceramics deriving from the metalloids silicon and carbon) bonded with a metal alloy. This definition allows cermets to be classed between metal matrix composites (see *Metal Matrix Composites: Matrices and Processing**) and ceramic matrix composites (see *Ceramic Matrix Composites: Matrices and Processing**). In both ceramic and metal matrix composites, a ceramic phase is added as reinforcement to a metallic or ceramic base material. In cermets, the metallic phase is, in most cases, added in order to improve toughness while maintaining ceramic-like properties: cermets are generally hard, refractory, and chemically resistant.

1. Classification

As in the present definition, cermets contain a dominant ceramic phase; these materials can be classified according to their refractory component, even if some of the properties of the composite may be quite different from those of the ceramic.

1.1 Hardmetals

When the term "hardmetal" was introduced (Schwarzkopf and Kieffer 1953), it was attributed to ceramics constituted of carbides, nitrides, borides, and silicides of the metals in the fourth to sixth groups of the periodic table of elements. Particular characteristics of these ceramics are relatively high electrical and thermal conductivities due to the metallic character of the chemical bonds. Their mechanical properties, such as hardness and brittleness, are typical of ceramics. In practice, most materials of commercial use that contain hardmetal ceramics are made with a metal binder, and the term "hardmetal" is now generally used to designate cermets based on these hardmetal ceramics.

The hardmetals that have had by far the largest commercial success are those used for the manufacturing of cutting tools. These materials are essentially based on the combination of tungsten carbide (WC), titanium carbide (TiC), or titanium carbonitride (TiCN), which constitute the majority of the ceramic phase, with an alloy dominated by one or more elements of the iron group: iron, nickel, cobalt. In the machining industry, the hardmetals that contain a dominant WC phase are generally named "cemented carbides" since their discovery by Schröter in 1922. The designation "cermet" has been historically reserved for TiC-based hardmetals, to specify the ceramic character (lower electrical conductivity) of TiC in comparison with WC. Among cemented carbides, those based on the system WC–Co are the most widespread, while the most used system among TiC-based cermets is TiC_xN_{1-x}–Mo–Ni.

The success of WC- and TiC-based hardmetals is due to their exceptional mechanical properties and to the favorable thermodynamic conditions in the processing of these materials by liquid-phase sintering. Carbides and nitrides of the hardmetals group are more stable than carbides and nitrides of the ferrous metals. Therefore, during sintering, the basic constituents of the cermet are not modified (apart from some potential exchange of metal ions between carbides and nitrides). The material thus behaves essentially like a binary alloy (Holleck and Kleykamp 1981) composed of well-defined metallic and ceramic phases. Moreover, partial dissolution of these carbides or nitrides into the metal produces a low melting point eutectic having a high solubility of the hard phase in the liquid metal. These conditions are very favorable for liquid-phase sintering (see Sect. 3).

Many other cermets based on hardmetal ceramics have been developed and tested for high-temperature applications, as wear parts, and in chemically aggressive environments. Their advantage compared to cutting tool hardmetals, which are also widely applied to wear parts, is often restricted, however, to special applications. A description of the main systems involving borides and silicides together with their industrial applications is given in Sect. 3.

1.2 Other Relevant Carbides and Nitrides

The hardest ceramics are found among nitride and carbide compounds of metalloid elements. Silicon carbide (SiC), silicon nitride (Si_3N_4), boron carbide (B_4C), cubic boron nitride (cBN), and diamond are the most important among these. Their use is, however, practically limited to metal matrix composites where these ceramics provide reinforcement or fulfill an abrasive function. The differences of elastic modulus and thermal expansion from metals, combined with a weakening of the ceramic grain boundary adhesion by products of reaction with metals, often make cermets based on these hard ceramics of little practical interest. Another compound that should be added to this class for its chemical characteristics is AlN. Due to good wetting by molten metals such as aluminum or nickel, cermets based on AlN can be produced by liquid-phase sintering. AlN is a good electrical insulator but has a very high thermal conductivity. Cermets based on the Al–AlN system are used in integrated circuits.

1.3 Oxide Cermets

Oxides used in industrial ceramics, such as Al_2O_3, ZrO_2, and SiO_2, are poorly wetted by most metals, with which they also tend to react at high temperature to form complex oxides. The processing of oxide cermets is therefore somewhat difficult (see Sect. 2). The high resistance to chemicals and the refractory character of oxide ceramics has, however, stimulated research efforts towards the engineering of cermets mainly aimed at improving the low mechanical and thermal shock resistance of the ceramic oxide.

2. Processing

Most cermet fabrication processes are based on powder metallurgy (PM) techniques. The metal and ceramic powders are mixed and then milled in ball mills or attritors. During the milling operation, a wetting organic lubricant is added in order to agglomerate the powder, prevent oxidation, and lubricate particle sliding. Organic lubricants may be polyglycols or paraffin wax dissolved in hexane, acetone, or alcohols. The purpose of the milling step is to obtain a homogeneous distribution of the ceramic grain size and to cover the ceramic powders with the metal, which facilitates the sintering. The slurry obtained after milling is then dried to remove the solvent. Spray drying is the most used technique for this purpose. After this operation the powder takes the aspect of spherical granules with a typical diameter of 0.1–0.5 mm.

Forming of the parts is achieved by powder compaction. The most used method for relatively simple shapes is cold pressing. Cold isostatic pressing, PM injection molding, or slip casting can be used for complex shapes.

After the cold processing steps, fully dense cermets are obtained by high-temperature PM processes. The most used method is sintering. This is generally carried out in batch ovens under vacuum, inert gas, or hydrogen. The choice between solid- and liquid-phase sintering depends on the particular ceramic–metal system. Liquid-phase sintering is well adapted to the systems used for hardmetal cutting tools described in Sect. 1.1. The densification, enhanced by the presence of a liquid phase, involves three stages: rearrangement of particles, solution-precipitation, and grain coarsening (German 1996). Rearrangement requires good wetting so that the capillary forces can promote the collapse of the powder structure. Solution-precipitation requires adequate solubility of the hard phase in the liquid phase. The ceramic grain coarsening must be controlled so as to achieve a homogeneous structure. All these requirements are generally met in the hardmetal families used for cutting tools. A typical sintering cycle for these hardmetals involves two or three steps where temperature, pressure, and sintering atmosphere are changed. First, the remainder of the organic lubricant is removed during the sintering cycle at 700–800 °C under vacuum. The true sintering step occurs at temperatures above the liquidus temperature of the ceramic–metal system. Typical temperatures for hardmetals are between 1350 °C and 1500 °C. The sintering step may be followed by hot isostatic pressing (HIP), which allows for elimination of residual porosity. This operation is efficient only on closed pores, which requires the achievement of a density greater than 95% after the pressureless sintering step. Solid-state sintering processes, such as warm extrusion and powder rolling, are used to form wires or slabs.

During the sintering process the microstructure, the morphology, and the phase partition of the cermets are built. In hardmetals, the ceramic phase composed of the carbides and/or nitrides partially dissolves in the binder and reprecipitates, producing the final shape and structure of the ceramic grains. In cemented carbides, the growth anisotropy of WC produces well-faceted single crystals (Fig. 1(a)) while most hardmetal carbides and nitrides like TiC have a NaCl cubic structure and produce round-shaped particles (Fig. 1(b)). The solubility of the other metallic elements in WC being very small, WC appears always pure, whatever the ceramic composition. However, the case of WC is special: hardmetal carbides, nitrides, and borides have a high reciprocal solubility and may exchange metal ions or carbon with nitrogen. Figure 1(b) shows, for example, a cermet sintered from TiCN, Mo_2C, and nickel powders. The ceramic grains assume a typical core–shell morphology. The core corresponds to the undissolved particles, while the shell is produced upon sintering by reaction and precipitation from the liquid phase of cubic Ti(Mo)C (Andrén *et al.* 1994). The shell and the core have identical crystal orientations.

Upon sintering, the ceramic particles join together in order to minimize the surface energy (German 1996) producing a continuous network. Depending on the binder content and on the dihedral angle (Wray 1976), the binder can also form an interconnected network. The connectivity of these two networks greatly influences the properties of the composite. In particular, the metal binder, which is in a constrained state in the network, exhibits features that are very different from those of the same but free metal. One example is that of cobalt, which in the hardmetals does not transform from the f.c.c. to the hexagonal structure (Exner 1979). In the metastable constrained state at room temperature, it can exhibit a highly strain-hardened structure with hardness comparable to that of the ceramic (Sigl and Fischmeister 1988); however, when constraints are released, e.g., by fracture, the cobalt can deform with a high toughening effect. The mechanical properties of the metal binder are also affected by solute impurities or precipitates deriving from the elements contained in the original ceramic powders.

Liquid-phase sintering is the most effective and economical production process when good wetting and solubility of the hard phase in the liquid are obtained; however, these conditions are not met in most metal–ceramic couples. Different techniques of pressure-assisted sintering, such as hot uniaxial or isostatic pressing, may then be adopted in order to obtain poreless materials in nonwetting systems like oxide cermets. These techniques may be applied also at relatively low temperatures in order to prevent undesired reactions between phases. Another processing technique that may be applied to cermets is infiltration of a ceramic matrix by molten metal. The excessive shrinkage that is associated with liquid-phase sintering may be avoided by this technique and accurate dimensions in complex shapes may be obtained. Infiltration is effective in wetting systems where the metal alloy can penetrate the preform by capillary forces. Pressure-assisted infiltration, which is widely used in metal matrix composite fabrication, can be used in nonwetting systems; however, it may be more difficult to apply this process in cermets due to the high volume content of the ceramic matrix.

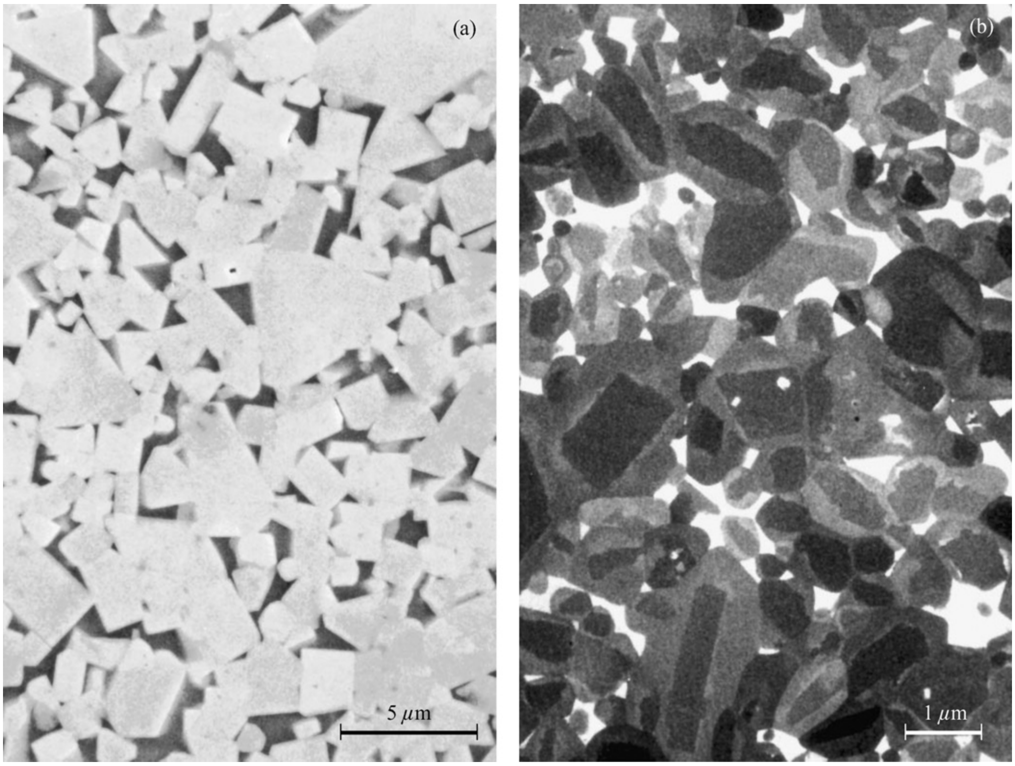


Figure 1
Grain structure of (a) WC–18 vol.% Co cemented carbide and (b) TiCN–6.2 vol.% Mo_2C–6.3 vol.% Ni cermet where typical core–rim structure can be observed.

Infiltration may also be applied to obtain graded cermets (see *Functionally Graded Materials**). Reactive synthesis techniques may be particularly attractive in making cermets. In most cases the sintering is achieved by an exothermic reaction between constituent powders. For example, TiAl–TiB_2 cermets have been produced by igniting a blend of titanium, aluminum, and boron (see XD process). Infiltration may also be associated with reactive synthesis. Systems like Al_2O_3–SiC–Al or ZrB_2–ZrC–Zr have been produced by infiltrating a carbide preform with a reactive metal in an oxidizing atmosphere.

3. Properties and Applications

3.1 Hardmetals for Cutting and Wear Applications

The improvement of the mechanical properties that can be obtained by combining metal with ceramic is particularly effective in the hardmetals for cutting tools, where an optimum combination of hardness and toughness is desired. From the commercial point of view, cutting tools represent the most widespread application of cermets in general. It is estimated that 65% of the total world production of tungsten is devoted to the fabrication of cemented carbides for cutting tools.

Since the beginning of cutting tool manufacturing, launched by the Krupp company in 1927 under the trademark Widia (wie Diamant), the WC–Co alloys have shown exceptional properties in cutting aluminum and cast iron. Their properties have since been enhanced successively with TiC and TaC additions to improve their resistance to oxidation and to retard diffusion into the tool edge (from elements like iron or nickel while machining steels), and with VC or Cr_2C_3 added to control the ceramic grain growth during sintering. Cemented carbides are still unsurpassed in room-temperature mechanical properties, showing a toughness (K_{IC}) of 10–24 MPa $m^{1/2}$ and corresponding Vickers hardness (roughly inversely proportional to toughness) in the range 2000–1100 H_v30 (Warren and Johannesson 1984). However, many cutting applications require a high working temperature; hence toughness and hardness are not the only parameters that define the quality of cutting tool materials. Their mechanical behavior at high temperature plays a fundamental role in defining their domain of application (Mari *et al.* 1999).

Generally, the softening of the metal phase with increasing temperature results in a decrease of the yield stress of the material. As the ceramic matrix constitutes the dominant phase in the material and forms a continuously bound skeleton, the deformation of the metal phase cannot produce extended cermet strain until the ceramic skeleton deforms. Plastic deformation of the ceramic grains is generally rather small, even in cubic carbides and nitrides where dislocations can glide in several high-density planes, as in f.c.c. metals. Above 1000 °C, grain boundary sliding of the carbides may however occur, leading to extensive plastic deformation. This phenomenon is particularly evident in cobalt-bonded cemented carbides. For this reason the TiC cermets, mostly used in the TiCN–Mo–Ni base compositions, find an important place in high-speed cutting applications where working temperatures can go beyond 1000 °C. Even if, just looking at the hardness and the toughness range ($H_v = 1400–1750$, $K_{IC} = 7–12$ MPa $m^{1/2}$ (Warren and Johannesson 1984)), these materials seem to perform less well than WC–Co, their refractory character proves extremely important to prevent wear while cutting.

It must be mentioned that in order to improve their chemical resistance to oxidation and to diffusion of external elements, most cutting tools are coated with refractory ceramics. The coating consists of a multilayered structure composed of TiC, Al_2O_3, and an external layer of TiN that gives a characteristic golden color to the cutting inserts.

3.2 Oxide Cermets

The development of oxide cermets is still limited despite a consistent research effort. The poor resistance of oxides to thermal shock and fracture could be improved by the presence of a tough metal binder; however, when a dominant oxide phase is present, the toughening effect of the metal is often too low. This is due, in many cases, to poor bonding at the metal/oxide interfaces. However, oxide cermets show some clear advantages when oxidation or chemical resistance are required, and also for high-temperature applications.

Al_2O_3 is used in combination with nickel, Ni_3Al, or chromium binders, or with TiC and a nickel alloy binder. These cermets are used for wear parts working at high temperature. Zr_2O–Ti or ZrO_2–Mo composites are used for wear and chemical applications. Its high thermal shock resistance and low thermal conductivity make ZrO_2 a good candidate as an insulator. ZrO_2–Mo cermets are used in thermocouple sheaths and extrusion dies in steel foundry technology. An interesting application of ZrO_2 in cermets is in the fabrication of anodes in solid oxide fuel cells made of yttria-stabilized ZrO_2–Ni cermets. These electrodes must be gas permeable, so the cermets contain 30% porosity. They are made by sintering a NiO–ZrO_2 ceramic in which the NiO is directly reduced during the first operation of the fuel cell. Another interesting application is the toughening of high T_c oxide superconductors (e.g., $YBa_2Cu_3O_7$) with a silver binder. These cermets are either hot extruded or made by reaction synthesis using Ag_2O as a silver precursor.

3.3 Borides

Borides have better chemical stability than carbides and may be used in applications that require extreme heat and corrosion resistance. Their brittleness, with strength hardly above 200 MPa makes their toughening with a metal binder very attractive. For example, adding 8% nickel to a TiB_2 matrix produces an increase of the flexural strength from 200 MPa to 700 MPa. TiB_2, ZrB_2, CrB_2, and MoB_2 combined with iron, nickel, chromium, or molybdenum binders are used as dies or crucibles with reactive metals like copper, bronze, lead, and zinc.

3.4 Silicides

Metallic silicides have found little commercial use due to their extreme brittleness. $MoSi_2$ is mostly used in heating elements due to its outstanding oxidation resistance combined with a favorable electrical resistance. $MoSi_2$-based cermets made with small amounts of nickel, cobalt, or silver as binder show a certain improvement of the toughness compared to the pure ceramic.

4. Summary

Cermets are constituted of a ceramic matrix bonded by a metallic binder. The metal improves the toughness and the thermal shock resistance of the ceramic. Two major groups of cermets can be defined: the hardmetals, which are based on carbides, nitrides, borides, and silicides of the fourth to sixth element groups, and the oxide cermets. Among hardmetals, the cemented carbides based on WC–Co–Ni and the TiC–Co–Ni cermets are used for the fabrication of cutting tools; these materials constitute the large majority of commercial cermets principally because they feature an optimum compromise between toughness and hardness. Cermets are mainly used in wear components, for chemical-resistant and high-temperature applications, and as insulators.

Bibliography

Andrén H-O, Rolander U, Lindahl P 1994 Phase composition in cemented carbides and cermets. *Int. J. Refract. Met. Hard Mater.* **12**, 107–13

Exner H E 1979 Physical and chemical nature of cemented carbides. *Int. Met. Rev.* **4**, 149–73

German R M 1996 *Sintering Theory and Practice*. Wiley, New York

Holleck H, Kleykamp H 1981 In: Hauser H H, Antes H W, Smith G D (eds.) *The Constitution of Cemented Carbide Systems*. Metal Powder Industries Federation, Princeton, NJ

Mari D, Bolognini S, Feusier G, Viatte T, Benoit W 1999 Experimental strategy to study the mechanical behaviour of hardmetals for cutting tools. *Int. J. Refract. Met. Hard Mater*. **17**, 209–25

Schwarzkopf P, Kieffer R 1953 *Refractory Hard Metals*. Macmillan, New York

Sigl L S, Fischmeister H F 1988 On the fracture toughness of cemented carbides. *Acta Metall*. **36**, 887–97

Warren R, Johannesson B 1984 In: Upadhyaya G S (ed.) *The Fracture Toughness of Hardmetals*. Elsevier, Amsterdam

Wray P J 1976 The geometry of two-phase aggregates in which the shape of the second phase is determined by its dihedral angle. *Acta Metall*. **24**, 125–35

D. Mari
Swiss Federal Institute of Technology, Lausanne, Switzerland

Clay-based Polymer Nanocomposites

A clay comprises silicate layers having a 1 nm thick planar structure. It has been shown that the silicate layers can be dispersed at the molecular level (nanometer level) in a polymer matrix with the polymer existing between the silicate layers (two-dimensional space). Polymer and clay nanocomposite materials produced in this manner are called polymer–clay hybrid materials. In this article some materials of this type are described, in categories based on synthetic methods. Five methods of synthesis of clay-based polymer nanocomposites will be described on the basis of the relationship between an ammonium cation-exchanged clay and monomer (or polymer). The methods are: (i) monomer intercalation, (ii) monomer modification, (iii) co-vulcanization, (iv) common solvent, and (v) polymer intercalation.

1. Classification of Producing Polymer–Clay Hybrid Material According to the Synthetic Method Employed

1.1 Monomer Intercalation Method

A polymerization producing nylon 6 is the ring-opening polymerization of ε-caprolactam. It can occur in the presence of clay after ε-caprolactam is intercalated into clay galleries, the silicate layers being dispersed uniformly in the nylon 6 matrix. We found that organophilic clay which had been ion-exchanged with 12-aminododecanoic acid was swollen by molten ε-caprolactam (basal spacing being expanded from 1.7 nm to 3.5 nm) (Usuki *et al.* 1993a). In the clay galleries the ε-caprolactam was polymerized, and the silicate layers were dispersed in the nylon 6 to give a nylon 6 clay hybrid (NCH) (Usuki *et al.* 1993b). This is the first example of an industrial clay-based polymer nanocomposite. Figure 1 shows a schematic representation of the polymerization.

The modulus of NCH increased to 1.5 times that of nylon 6, the heat distortion temperature increased to 140 °C from 65 °C, and the gas barrier effect was doubled at a low loading (2 wt.%) of clay (Kojima *et al.* 1993b).

There is another example, in which ε-caprolacton is polymerized in the clay galleries in the same manner. In this case, the gas permeability decreased to about 20% under 4.8 vol.% (12 wt.%) clay addition (Messersmith and Giannelis 1995). There is yet another example of an epoxy resin clay nanocomposite. In this case, the tensile strength and modulus

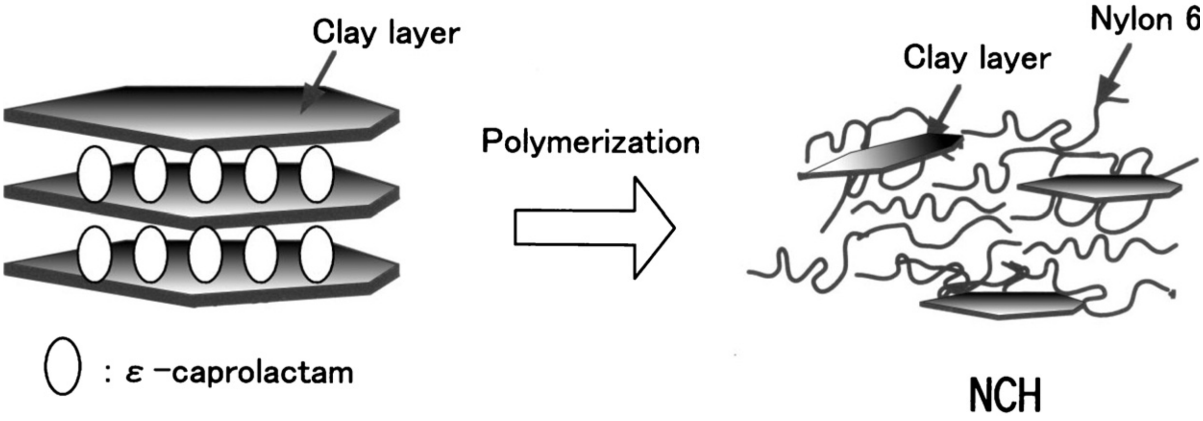


Figure 1
Schematic diagram of polymerization to NCH.

increased drastically on 2–20 wt.% clay addition (Lan and Pinnavaia 1994).

1.2 Monomer Modification Method

In one example, a quaternary ammonium salt of dimethylaminoacrylamide (Q) was ion-bonded to silicate layers, and ethyl acrylate (EA) and acrylic acid (Aa) were copolymerized in the clay galleries. The ratio between EA and Aa was 10:1 (mole ratio). Each of the four kinds of acrylic resin–clay hybrid were polymerized. The clay content was 1, 3, 5, and 8 wt.% on the basis of solid acrylic resin. A suspension above 3 wt.% clay addition acted as pseudoplastic fluid (Tables 1 and 2 give details of compositions). The acrylic resin clay hybrid films cross-linked by melamine were transparent, and the gas permeability of their film (Fig. 2) decreased to about 50% on 3 wt.% clay addition (Usuki *et al.* 1995).

There are also some other reports of acrylic resin–clay hybrids. A poly(methyl methacrylate) clay hybrid was synthesized using a modified organophilic clay in the same manner (Biasci *et al.* 1994), and by emulsion polymerization (Choo and Jang 1996). Figure 3 shows a schematic representation of this polymerization method.

1.3 Co-vulcanization

The basal spacing of organophilic clay ion-bonding nitrile rubber (NBR) oligomer with telechelic amino groups was expanded 0.5 nm from its initial spacing (1.0 nm) (Moet *et al.* 1994). After this, high molecular weight NBR was kneaded with this organophilic clay and vulcanized with sulfur. It produced an NBR clay hybrid consisting of dispersed clay and co-vulcanized high molecular weight NBR and NBR oligomer (Fukumori *et al.* 1991). Its permeability to hydrogen and water decreased to 70% on adding 3.9 vol.% clay (Kojima *et al.* 1993a). Figure 4 shows a schematic representation of this production method.

1.4 Common Solvent Method

In the case of synthesis of polyimides, the polymerization solvent for the polyamic acid precursor of polyimide is usually dimethyl acetamide (DMAC). We found that a clay ion-exchanged dodecyl ammonium ion could be dispersed in DMAC homogeneously. The solution of this organophilic clay and DMAC was added to a DMAC solution of polyamic acid. The film was cast from a homogeneous mixture of clay and polyamic acid, and was heated at 300 °C to get the desired polyimide clay hybrid film. The permeability of water decreased to 50% on adding 2.0 wt.% clay (Yano *et al.* 1993). It was confirmed that the permeability of carbon dioxide also decreased by half (Lan *et al.* 1994). Figure 5 shows a schematic representation of this method.

1.5 Polymer Intercalation Method

Polypropylene (PP) clay hybrids could not be synthesized easily because PP is hydrophobic and has

Table 1
Composition of acrylic resin.[a]

Acrylic resin	Ethyl acrylate, EA (g)	Acrylic acid, Aa (g)	Dimethylamino acrylamide, Q (g)	Butyl cellosolve (g)	*NV, calculated*[b] (%)
A	144.7 (90.8)	10.4 (9.08)	0.51 (0.12)	157.7	46.0
B	144.7 (90.6)	10.4 (9.04)	1.60 (0.36)	157.7	46.1
C	144.7 (90.4)	10.4 (9.04)	2.54 (0.56)	157.7	46.2
D	144.7 (90.1)	10.4 (9.01)	4.13 (0.89)	157.7	45.0

[a] Figures in parentheses are mole ratios. [b] NV: nonvolatile concentration.

Table 2
Composition of acrylic resin–clay hybrid.

Acrylic resin–clay hybrid	Acrylic resin (g)	NV, measured[a] (%)	Deionized water (g)	Montmorillonite–water suspension, 4 wt.% (g)	*NV* in montmorillonite[a] (%)
A	a 14.0	44.9	16.6	1.6	1.0
B	b 14.0	45.2	11.6	4.9	3.1
C	c 13.4	45.3	10.5	7.5	4.9
D	d 14.5	43.9	6.4	12.7	8.0

[a] NV: nonvolatile concentration.

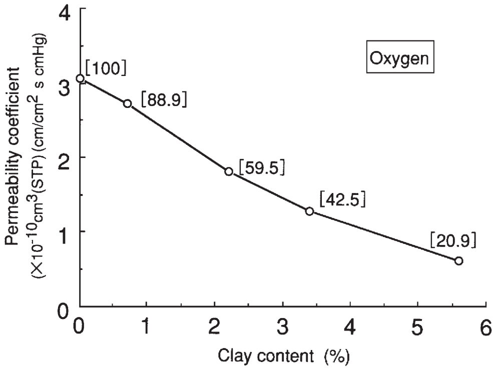


Figure 2
Permeability coefficient of acrylic resin clay hybrid film.

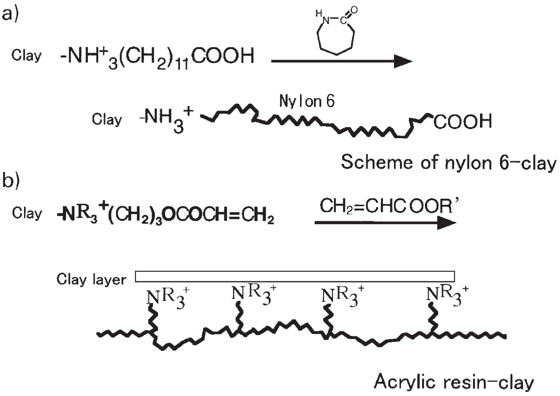

(a)
Clay -NH⁺₃(CH₂)₁₁COOH
Clay -NH₃⁺
Nylon 6
COOH
Scheme of nylon 6–clay
(b)
Clay -NR₃⁺(CH₂)₃OCOCH=CH₂
CH₂=CHCOOR'
Clay layer
NR₃⁺
Acrylic resin–clay

Figure 3
Scheme of NCH and acrylic resin clay hybrid.

poor miscibility with clay silicates. Dioctadecyldimethyl ammonium ion was used as a modifier for the clay and a polyolefin oligomer was used so that the clay became more compatible. Organophilic clay, polyolefin oligomer, and PP were blended using an extruder at 200 °C. It was confirmed by transmission electron microscopy that the clay was dispersed in a monolayer state in the PP matrix. Thus PP was directly intercalated into the clay gallery (Usuki *et al.* 1997). Figure 6 shows a schematic representation of this polymerization.

There is also a direct-intercalation process in which PP is modified using maleic anhydride, followed by melt compounding (Kawasumi *et al.* 1997). It is a useful process from an industrial standpoint. There are also some studies showing that polymer is intercalated directly into the clay galleries. There is a report that intercalation of nylon into clay galleries was successful (Takemura *et al.* 1996). The clay silicates may not have been dispersed uniformly, judging by the physical properties exhibited.

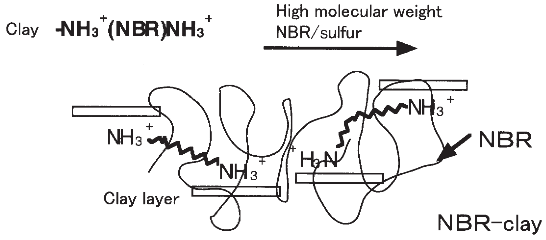

Clay -NH₃⁺(NBR)NH₃⁺
High molecular weight
NBR/sulfur
NH₃⁺
NH₃⁺
H₂N
NH₃⁺
NBR
Clay layer
NBR–clay

Figure 4
Scheme of NBR clay hybrid.

Scheme of NBR clay hybrid.

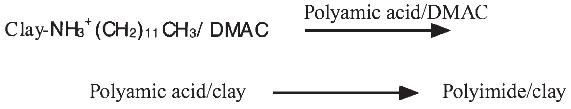

Clay-NH₃⁺(CH₂)₁₁CH₃/ DMAC
Polyamic acid/DMAC
Polyamic acid/clay
Polyimide/clay

Figure 5
Scheme of polyimide clay hybrid.

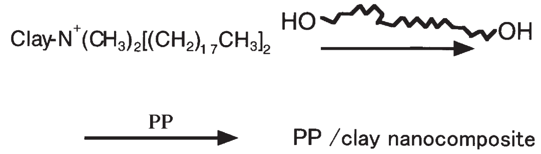

Clay-N⁺(CH₃)₂[(CH₂)₁₇CH₃]₂
HO
OH
PP
PP /clay nanocomposite

Figure 6
Scheme of PP clay hybrid.

2. *Functional Clay Hybrids*

The examples described above involved organic polymer clay hybrids which are aimed at high performance structural polymer materials with high strength, high modulus, and high gas barrier properties. This hybrid technique should also be applicable to various other types of materials, including low molar mass compounds. This section presents one of the unique attempts to obtain novel hybrid materials based on a low molar mass liquid crystal and organized clay mineral (liquid crystal clay composite, LCC). Clay exchanged with 4-cyano-(4′-biphenyloxy)undecyl ammonium ion was used to enhance the miscibility between clay and the liquid crystal. The LCC (clay content 1.25 wt.%) cell displayed a reversible and bistable electro-optical effect based on light scattering which could be controlled by changing the frequency of an electric field (Kawasumi *et al.* 1996). Figure 7 shows a typical change in the light transmittance of the LCC (clay content 1.25 wt.%) when the cells were subjected to electric fields.

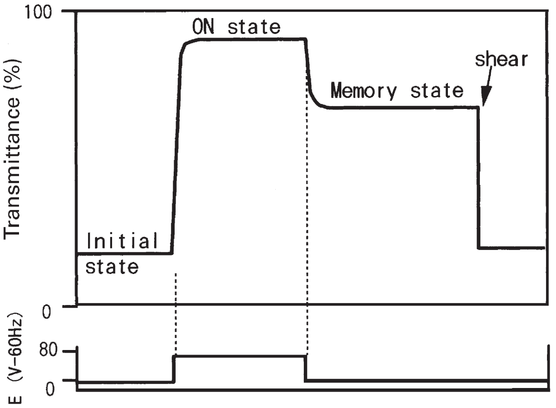


Figure 7
Light transmittance of LCC.

This new material would be a potential candidate for advanced applications such as a light controlling glasses, high information display devices which do not require active addressing devices, erasable optical storage devices, and so on.

3. *Possible Future Subjects*

The above examples illustrate the synthesis methods and the properties of clay-based nanocomposites including polymers and liquid crystals. A modification of the clay surface plays a principal part in the technology. Other modifications of monomer and/or polymer, or an addition of adequate compatibilizer for organic materials and clay could open a new methodology of clay nanocomposite production. Future subjects are considered to be:

(i) Clay-based nanocomposite technology involving polymer alloys.

(ii) Investigations of clay-based nanocomposites involving new properties such as transparency, recyclability, fire retardation (Gilman and Kashiwagi 1997), and biodegradation.

(iii) Identification of the most suitable clay manufacturing technique for each polymer.

Bibliography

Biasci L, Aglietto M, Ruggeri G, Ciardelli F 1994 Functionalization of montmorillonite by methyl methacrylate polymers containing side-chain ammonium cations. *Polymer* **35**, 3296–304

Choo D, Jang L W 1996 Preparation and characterization of PMMA–clay hybrid composite by emulsion polymerization. *J. Appl. Polym. Sci.* **61**, 1117–22

Fukumori K, Usuki A, Sato N, Okada A, Kurauchi T 1991 Rubber–clay molecular composites: nitrile rubber/layered silicate system. In: *Proc. 2nd Japan Int. SAMPE Symp.* SAMPE, Covina, CA, pp. 89–96

Gilman J W, Kashiwagi T 1997 Nanocomposites: a revolutionary new flame retardant approach. In: ADDITIVES '97 Executive Conference Management

Kawasumi M, Hasegawa N, Kato M, Usuki A, Okada A 1997 Preparation and mechanical properties of polypropylene–clay hybrids. *Macromolecules* **30**, 6333–8

Kawasumi M, Usuki A, Okada A, Kurauchi T 1996 Liquid crystalline composite based on a clay mineral. *Mol. Cryst. Liq. Cryst.* **281**, 91–103

Kojima Y, Fukumori K, Usuki A, Okada A, Kurauchi T 1993a Gas permeabilities in rubber–clay hybrid. *J. Mater. Sci. Lett.* **12**, 889–90

Kojima Y, Usuki A, Kawasumi M, Okada A, Fukushima Y, Kurauchi T, Kamigaito O 1993b Mechanical properties of nylon 6–clay hybrid. *J. Mater. Res.* **8**, 1185–9

Lan T, Kaviratna P D, Pinnavaia T J 1994 On the nature of polyimide–clay hybrid composites. *Chem. Mater.* **6**, 573–5

Lan T, Pinnavaia T J 1994 Clay-reinforced epoxy nanocomposites. *Chem. Mater.* **6**, 2216–9

Messersmith P B, Giannelis E P 1995 Synthesis and barrier properties of poly(ε-caprolacton)–layered silicate nanocomposites. *J. Polym. Sci., Polym. Chem.* **33**, 1047–57

Moet A, Akelah A, Salahuddin N, Hiltner A, Baer E 1994 Layered silicate/ATBN nanocomposites. In: Gonsalves K E, Chow G, Xiao T D, Cammarata R C (eds.) *Mat. Res. Soc. Symp. Proc.* Materials Research Society, pp. 163–70

Takemura K, Hosokawa T, Tamura K, Iniue H 1996 Polymer–clay composites. *Japan–China Seminar on Advanced Engineering Plastics, Polymers, Alloys and Composites.* Society of Polymer Science, pp. 32–5

Usuki A, Kato M, Okada A, Kurauchi T 1997 Synthesis of polypropylene–clay hybrid. *J. Appl. Polym. Sci.* **63**, 137–9

Usuki A, Kawasumi M, Kojima Y, Okada A, Kurauchi T, Kamigaito O 1993a Swelling behavior of montmorillonite cation-exchanged for ω-amino acids by ε-caprolactam. *J. Mater. Res.* **8**, 1174–8

Usuki A, Kojima Y, Kawasumi M, Okada A, Fukushima Y, Kurauchi T, Kamigaito O 1993b Synthesis of nylon 6–clay hybrid. *J. Mater. Res.* **8**, 1179–84

Usuki A, Okamoto K, Okada A, Kurauchi T 1995 Synthesis and properties of acrylic resin–clay hybrid. *Kobunsi Ronbunshu* **52**, 728–33

Yano K, Usuki A, Okada A, Kurauchi T, Kamigaito O 1993 Synthesis and properties of polyimide–clay hybrid. *J. Polym. Sci., Polym. Chem.* **31**, 2493–8

A. Usuki
Toyota Central R&D Labs, Aichi, Japan

Composite Dental Materials: Wear

The use of light-cured composite resins for restoring posterior stress-bearing cavities has increased significantly in recent years. These materials feature the advantage of good esthetics for posterior restorations and are able to bond to tooth structures (see Bayne *et al.* 1989, Lutz *et al.* 1984, Manhart *et al.* 2000, Roulet 1988). New methods for diagnosis and treatment, such as Diagnodent (KaVo) and caries detectors (Cavex), Sonicsys (KaVo), micropreparation

(Komet), and air abrasion have ensured that composite resins are the materials of choice in the concept of minimal invasive dentistry (Degrange and Roulet 1997). As an alternative to amalgam, the early attempts to place composites in posterior teeth had only limited success because of inadequate material properties. Inadequate resistance to wear, resulting in a loss of anatomic form under masticatory abrasion and attrition, fracture within the body of the restoration and margins, and marginal leakage due to polymerization shrinkage were often cited as being the most common causes of failure in posterior composite restorations (Manhart *et al.* 2000). Improvements in the properties of the materials together with their positive clinical performances encourage continuously the use of posterior composite resins as a viable alternative to amalgam (see Bayne *et al.* 1989, Degrange and Roulet 1997, Lutz *et al.* 1984, Manhart *et al.* 2000, Roulet 1988).

Because it is natural tooth tissues that are being replaced, human enamel and dentin should be used as the physiological standards with which to compare composite resins, especially in the posterior region. The goal of restorative dentistry should be to make highly wear-resistant occluding surfaces that do not cause wear on the opposing structures (Lambrechts *et al.* 1989, Lambrechts *et al.* 1991, Willems *et al.* 1993a).

1. Description

1.1 Mechanisms

Wear of dental composites includes diverse phenomena such as adhesion, abrasion, attrition, chemical degradation, and fatigue, to mention the most common (Lim *et al.* 2002). In reality, the various mechanisms usually act simultaneously and/or sequentially and often influence each other in a very complex way, resulting in the general tendency of wear processes to be very difficult to predict (Pallav 1996).

Abrasion in the dental literature is caused mainly by frictional surface interactions with toothbrush and paste, food bolus, and fluid components during chewing. This form of wear is considered an important mechanism of occlusal material loss with composites (Krejci *et al.* 1990). Attrition is caused by direct contact of sharp roughness asperities of the antagonist, which should at least be some 50% harder than the wearing substrate for substantial wear to occur. Attrition may cause substantial changes in surface texture (roughness, smear layer, etc.). Additional, chemical wear is an underestimated wear mechanism. All kinds of chemical influences on wear processes are possible, such as degradation of silane couplings in composites and possible softening of resin material by food or saliva constituents. Surface fatigue is a wear phenomenon that occurs in moving contacts and/or at cyclic load variations. Due to common dental geometry, stress concentrates slightly below the surface. Because of this, a network of subsurface cracks is created. One of the characteristics of surface fatigue is the incubation period (McKinney and Wu 1982, Lambrechts *et al.* 1987, Juvinall and Marshek 1991). The cracks are very likely to grow and connect at their roots, through stress concentration (Truong and Tyas 1988) until, at the end of the onset period, the damaged area no longer offers enough support for the superficial layer, which will then rapidly delaminate, leaving an extremely rough surface (Lambrechts *et al.* 1985).

Another important wear mechanism is adhesive wear. High adhesive shear stresses tear platelet-shaped wear particles out of the surface(s), which may, sometimes permanently, remain attached to the antagonistic surface, which in mild cases reduces the wear rate and in severe cases can increase the wear rate.

The wear process of dental composites is complicated by the fact that the filler and matrix consist of fundamentally different materials, the relative influence of which varies greatly with the wear type. With abrasion, the relatively soft matrix allows the digging out of filler particles, while with direct contact, attrition and adhesive wear are located at the peaks of protruding particles. If, however, the filler is also relatively soft, then the antagonist may be smoothed by material transfer and in this way, wear may be reduced (Pallav 1996).

1.2 Types of Wear

Wear of composite resins takes place at the occlusal surface and at the proximal interface. In literature, there is a difference between wear at contact free occlusal areas (CFOA) and wear at the occlusal contact area (OCA) (Lambrechts *et al.* 1987).

In the chewing cycle, opposing dentition traps a layer of food and grinds it as the teeth move past one another. The excursion of food particles produces wear of the tooth structure and restorative materials through three-body abrasion. This is considered the main wear mechanism active in the contact-free area, resulting in generalized loss of form (Lutz *et al.* 1984). The chewing forces produced during this phase have been modeled in the range of 10–20 N (De Gee *et al.* 1986, De Gee and Pallav 1994, Sakaguchi *et al.* 1986). At the end of the chewing cycle, the sliding motion stops as the teeth reach centric occlusion. The chewing force is heightened within the range of 50–150 N (Bates *et al.* 1975), and tooth surfaces come into more direct contact. This produces localized two-body abrasion wear or attrition at sites termed OCAs (Lutz *et al.* 1984). Wear on the OCAs was found to be approximately 2.5 times greater than that on the CFOAs (Roulet 1988).

Proximal surfaces are subjected to compressive and sliding forces during function as well. As a result, interproximal wear of posterior resin composite restorations can be significantly greater than that of unrestored surfaces (Cunningham *et al.* 1990, Wang *et al.* 1989). The periodontal status and its related tooth mobility determine the degree of sliding and related wear.

1.3 Influencing Factors

Generally, wear can be related to either material or clinical factors. Material factors relate to the resin composite's particle size, shape and hardness, the filler content, the interparticle spacing, the filler distribution, and degree of conversion. Additionally, the silane coupling, the nature of the matrix, and the surface hardness play an important role in the composite materials wear resistance (Hilton 2000, Suzuki *et al.* 1995, Venhoven *et al.* 1996). Since the process of three-body abrasion involves the cutting away of a soft material by a hard abrasive, it seems reasonable that adding hard inorganic filler particles to a soft resin matrix should enhance the material's overall resistance to abrasion (Leinfelder 1988). An important factor is particle size. In the range of current composites a smaller particle size is desirable. It appears that there is a critical value of filler particle size (1.2–1.5 μm), under which the food fibers are not able to penetrate the interparticle space, so the erosive capability of the erosive medium will be reduced (Venhoven *et al.* 1996). *In vitro* wear tests indicated that composites containing the smallest (0.2 μm) spherical particles exhibited substantially higher wear resistance compared to larger filler paricles (Suzuki *et al.* 1995). However, the mean particle size does not adequately describe the filler system. In particular, the proportion of the largest fillers dominates the wear behavior (Degrange and Roulet 1997). An additional aspect to be considered is particle brittleness. Under high stress, brittle particles will easily fracture and cause rapid abrasion of the material itself, but less antagonistic engagement (Lambrechts *et al.* 1991).

Besides the addition of filler particles to the matrix of the composite, filler–matrix-interactions appear to have great effect on increasing the wear resistance of the composite resins. Apparently, the ability for a resin composite to resist abrasive action is an interaction of the different mechanical factors. Clinical studies have shown less heavily filled composites ($<60\%$ filled by volume) to exhibit unacceptable wear (Willems *et al.* 1993b, 1993c). However, the results from clinical studies indicate that no direct relationship exists between filler content and abrasion resistance since wear tribology is multifactorial and complex. Also, *in vivo* wear measuring techniques are often unable to unravel the complicating biological factors. Because of their generally lower filler content, (30–50%), microfilled composites are more subjected to attrition and marginal breakdown, especially adjacent to OCAs (Manhart *et al.* 2000, Collins *et al.* 1998). However, they are more resistant to abrasion because of their smoother surface, decreased interparticle spacing, and decreased friction to food particles.

Bayne *et al.* (1992) suggested that, in a microfilled composite, very little (<6.0 vol.%) microfiller was required theoretically to provide matrix protection. Pallav *et al.* (1989) reported that an optimum microfiller level was approximately 15 vol.% to achieve the good wear resistance for small-particle hybrid composites, if the fillers were distributed well. As filler volume increased, wear was reduced regardless of filler treatment. Wear resistance of microfilled composites was significantly enhanced by filler volume with an increase from 25 vol.% to 30 vol.%.

Besides the filler type and filler volume fraction, the interparticle spacing and filler distribution influences the physical properties and wear behavior of dental composite materials (Degrange and Roulet 1997, Bayne *et al.* 1992, Mitchem and Gronas 1990, Ferracane and Condon 1999).

From Jørgensen's findings (Jørgensen and Asmusson 1978, Jørgensen *et al.* 1979), it may be concluded that the use of finer particles for a fixed volume fraction of filler results in decreased interparticle spacing and reduced wear. Filler particles situated very close protect the softer resin matrix from abrasives, thus reducing wear. Based on research (Jørgensen and Asmussen 1978, Jørgensen *et al.* 1979) and clinical research data available today, it appears that the critical interparticle spacing for dental composites is around 0.1–0.2 μm.

In addition to interparticle spacing, another key factor in wear of composites is good stress-transfer ability (Söderholm and Shang 1993, Calais and Söderholm 1998). Such a stress transfer can be enhanced by either improving the bond between the filler and the matrix, or by increasing the filler surface area (Degrange and Roulet 1997). Silane coupling agents are thought to play a major role in enhancing the adhesion of the interface between the inorganic filler and the organic resin. The silane used in most dental composites is γ-methacryloxypropyl-trimethoxysilane, which can bond to inorganic silica surfaces and organic polymers (Söderholm and Shang 1993). Another advantage of the use of a coupling agent in dental composites is that it, at least to some extent, protects the filler against hydrolytic degradation. Condon and Ferracane (1997) and Beatty *et al.* (1998) reported that wear resistance decreased linearly as the percentage of silane-treated silicate glass fillers in experimental hybrid composites was reduced.

There is an advantage gained by the reinforcement of the resin matrix with well-bonded microfillers, which protect the matrix and interfere with crack propagation at higher filler levels (Mohsen and Craig

1995). It is important to keep microfillers well dispersed to achieve the full effect of protection (Bayne *et al.* 1992). Particles in which silane molecules have condensed may form clusters that are not penetrated by the resin monomer.

Filler particle clustering ought to be one of the detrimental factors to performance of particle-reinforced composites. According to Söderholm (1985), the silane coupling agent tends to form aggregates on the filler surface. Aggregates of filler particles and silane coupling, however, do not form a very stable bond between filler and resin. Mohsen and Craig (1995) suggested that dispersion of the filler in the monomer mixtures involved wettability of the filler surface by the monomer, dispersion of the particles, and stabilization. The degree of agglomeration decreased slightly with an increase of the filler content. This may be a result of the increasing amount of filler that is available to break down the filler agglomerates by grinding against one another during mixing. Bayne *et al.* (1992) calculated the minimum and maximum interagglomerate spacing. They presumed that agglomerates were 0.5–1.0 μm in diameter with 1000 microfiller particles being involved, and particles in the agglomerate were assumed to be closely packed. To achieve the full effect of protection, they concluded that the critical filler fraction should be in the range of 35–48 vol.%. In addition, contamination of the filler might disturb silanization and reduce filler–matrix coupling. Therefore, Shirai *et al.* (2000) proposed the use of sodium peroxodisulfate in a presilanization filler decontamination step in the process for the production of resin composites (Yoshida *et al.* 2002, Shirai *et al.* 2000).

Wear was found to reduce significantly as the degree of cure was increased (Ferracane *et al.* 1997). The excellent correlation between degree of conversion of the polymer matrix and abrasive wear for experimental composites verifies that, for a given filler composition, abrasive wear is dependent upon the properties of the polymer network. It is expected that the increased cure produced by the longer curing times causes the polymer network to be more highly cross-linked, and, therefore, more resistant to the forces of abrasion. It should be noted that the baseline depths are also related to the curing time, because the composites with lower conversion are less abrasion-resistant and are polished to a greater depth during the finishing sequence. The heat-cure process enhances the fracture toughness of the microfill composite, which results in an enhanced resistance to marginal breakdown (Ferracane and Condon 1992). Study results imply that maximum wear resistance requires that the composite should be cured to its maximum amount (Ferracane *et al.* 1997).

However, in addition the strength of the resin matrix must be considered. A weak, incoherent matrix can invoke phenomena such as plucking out of filler particles and thus reduce the abrasive capacity. Studies have revealed that matrix composition is a significant factor affecting the wear rate. The superior resistance of UEDMA/TEGDMA resin can be related to its higher degree of conversion (Condon and Ferracane 1996, Söderholm and Shang 1993). In laboratory evaluations, the UEDMA/TEGDMA had a conversion rate of 70% while the BisGMA/TEGDMA had a conversion rate of 55% (Söderholm *et al.* 2001).

It is noteworthy that long-term exposure to water after 3 yr caused a DPMA-based matrix to undergo substantially greater degradation than an UDMA-based matrix (Beatty *et al.* 1993).

The BisGMA resin compound is susceptible to chemical softening, which reduces the surface hardness of the composites and enhances the occlusal wear rate (McKinney and Wu 1982). The mixing of certain incompatible resins such as BisGMA and UDMA can cause agglomeration if the resins are not stirred well (Lambrechts and Vanherle 1983). Thus, good mixing is very important to avoid filler agglomeration and an inhomogeneous resin phase.

The chemical environment is one aspect of the oral environment which could have an appreciable influence on the *in vivo* degradation of composite resins. The resin matrix can be softened and fillers can be leached out when composites are exposed to certain chemicals/food simulating liquids (Yap *et al.* 2002).

Apparently the toughness of the resin matrix is an important factor since urethane-based materials are the most abrasion resistant (Beatty *et al.* 1998). The intrinsic surface roughness of composite resins must be equal to or lower than the surface roughness of human enamel on enamel-to-enamel occlusal contact areas ($Ra = 0.64\,\mu m$) (Lambrechts *et al.* 1991). Roughness determines the biological strength of composite resins. The nanoindentation hardness value of the filler particles (2.91–8.84 GPa) must not be higher than that of the hydroxyapatite crystals of human enamel (3.39 GPa) (Willems *et al.* 1993a). However, an ultrafine filler size will reduce the impact of the filler hardness.

Whenever filler particles protrude and are large and extremely hard, there can be high antagonistic wear rates that would cause catastrophic loss of tooth substance in time (Lambrechts *et al.* 1987). Therefore, if there is a mismatch between the filler hardness and surface roughness of the composite and that of enamel, there is a danger that the enamel will be worn down. During mastication, composite filler particles may scratch and abrade the antagonistic enamel. The vertical dimension will thus decrease if premolar and molar regions are treated with such restorative material.

A smooth surface provides reduced frictional wear at the occlusal contact area. This smoothness will benefit the composite wear as well as the antagonistic enamel wear (Lambrechts *et al.* 1988).

Wear resistance tends to increase as cavity size decreases. The restorations placed should be as small as possible in order to benefit from the strong support of, and a full protection by, the surrounding natural tooth structure. It was also found that more wear occurred with larger-sized restorations (Roulet 1988).

Tribology or wear science plays an important role. The width of various parts of the restoration is of influence in three ways. First, at the central part of the fossa the erosive activity increases with the distance from the center. This implies that the average wear rate at a restored area in the central part increases with the width of the restoration. Second, the greater (wider) the restored part of a tooth's occlusal surface, the more occlusal contacts are, statistically, on the restored part, and the more a developing height loss will be compensated for by repositioning (natural eruption) of the teeth. Because of this, the wear rate will decrease less with developing height loss. Third, imagine a cusp and fossa with three equal escape ways. If one of its escape ways contains a wider restoration, then its flow resistance will decrease more with height loss than that of the other two. Because of this, gradually more food will escape at the wider restoration and less will flow out at the other two. For the height loss versus time means that, although the wear rate decreases at all the three escape ways, it will decrease less at the wider restoration and more at the other two (Pallav 1996).

Another factor determining the wear process is the dimension of the occlusal contact area. Krejci *et al.* (1991) found that composite wore significantly less with increased OCA dimensions for a constant load. When considering an OCA with restricted dimensions, the chewing stresses will cause more local composite damage because the occlusal forces will be spread over a tiny contact area. Increasing the OCA will diminish composite wear (Krejci *et al.* 1991, Willems *et al.* 1993c).

2. *Measuring Methods*

Several systems have been described for the assessment of changes in anatomical form of composites and major discrepancies have been shown between the *in vivo* and *in vitro* wear rates of posterior composites (Roulet *et al.* 1987). Attempts to correlate *in vitro* results with a long-term *in vivo* situation have not been very successful (Sakaguchi *et al.* 1986, Taylor *et al.* 1994). Also the complex *in vivo* wear behavior cannot be predicted from physical and mechanical testing alone. Therefore, long-term quantitative *in vivo* wear measurements are still essential to determine the clinical performance of new posterior composite materials. Clinical evaluations are generally the most widely accepted measures of composite performance. However, clinical trials are both expensive and time-consuming. In addition, the lack of control over important variables such as chewing force, dietary intake, or environmental factors in *in vivo* studies limits their contribution to the tribology of dental materials (Taylor *et al.* 1994). It is important that any machine which attempts to characterize composite wear behavior reproduces both of these types of wear—abrasion and attrition (Condon and Ferracane 1997). The OHSU oral wear simulator developed by Condon and Ferracane (1996) has demonstrated its ability to produce clinically relevant wear results. It has the ability to generate wear simultaneously through the two dominant oral wear mechanisms of abrasion and attrition, while also allowing for the measurement of opposing enamel wear attrition (Condon and Ferracane 1997). Other oral wear simulators are the Munich Artificial Mouth, the ACTA oral wear simulator, and the Artificial Mouth from Minnesota (Manhart *et al.* 2000, Mehl *et al.* 1997, Pallav 1996, Sakaguchi *et al.* 1986). Experience with a denture model has verified that similar abrasive wear of composites occurs as in dentate patients (Condon and Ferracane 1996, Mitchem and Gronas 1990, Söderholm 1985).

The results from *in vitro* and *in vivo* wear tests can be measured in several ways. *In vivo* wear can be evaluated in a direct way, using clinical performance scales like the US Public Health Service criteria (Cvar and Ryge 1971). Clinical wear performance can also be measured in an indirect way, using the same techniques as for *in vitro* wear measurements. The *in vivo* wear can be measured by the use of replicas (Lambrechts *et al.* 1991, Manhart *et al.* 2000, Söderholm *et al.* 2001). The most common measuring method is profilometry (Condon and Ferracane 1997, Ferracane and Condon 1999, Ferracane *et al.* 1997, Lutz *et al.* 1984, Söderholm *et al.* 1992). The distinct wear step between enamel and composite can be visualized by means of scanning electron photomicrographs (Willems *et al.* 1993b).

Other possible methods of determining the wear on replicas are the use of a 3D-measuring technique, a scanning microscope or a 3D-laser scanning measurement method (Lambrechts *et al.* 1984, Mehl *et al.* 1997, Söderholm *et al.* 1992, Yoshida *et al.* 2002). Figure 1 is an excellent example of wear measurement by means of a 3D-laser scanner.

An *in vitro* investigation found that there were no significant differences in consistency between the microscopic measure and the laser measure (Söderholm *et al.* 2001). It seems reasonable to suggest that the digital laser scanning method gives depth values comparable to microscopic evaluation, but because of its ability to register the entire surface area, this technique has certain advantages in clinical wear evaluations of dental materials over the microscopic evaluation method (Yoshida *et al.* 2002).

The mass of data generated about the various competing products is very confusing, and ranking the different materials according to their laboratory results does not necessarily reflect their clinical performance (Lambrechts *et al.* 1987). This is not

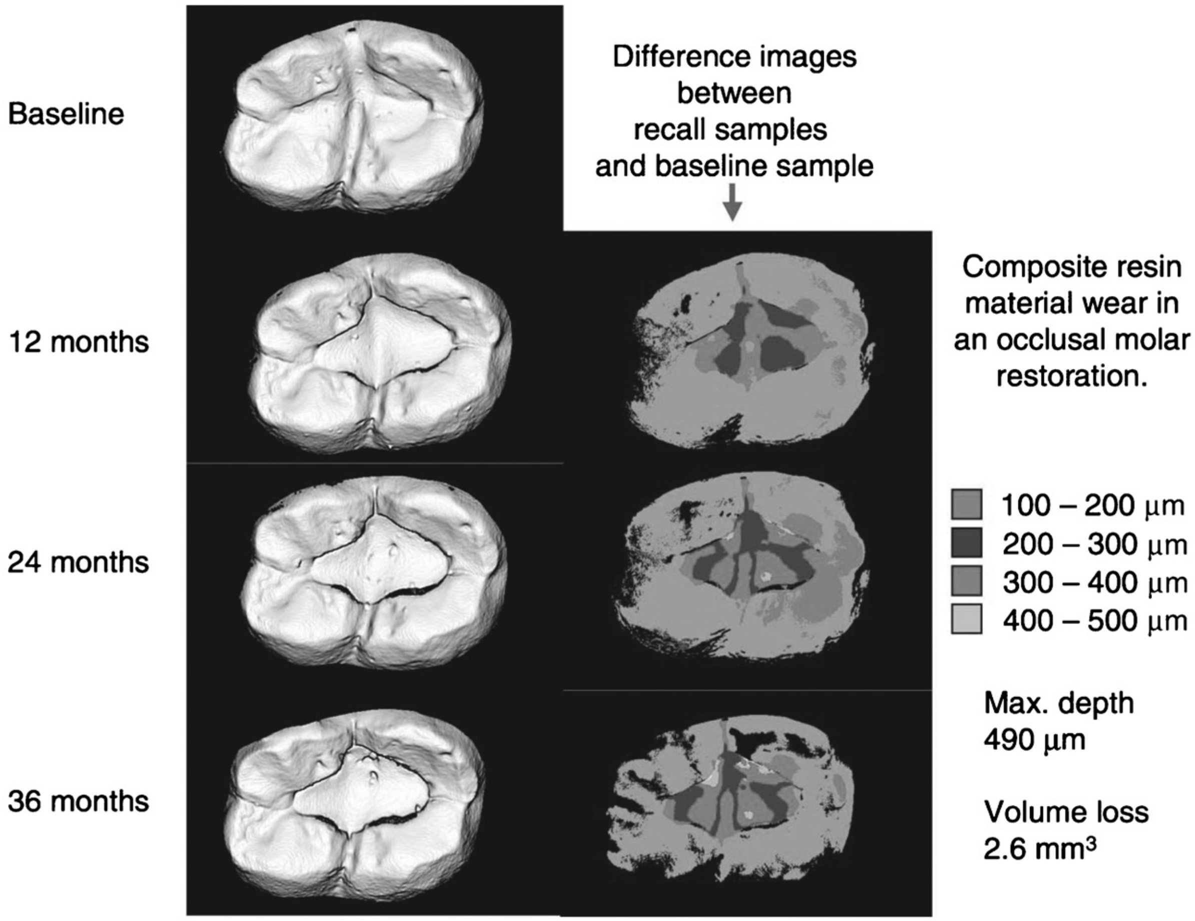

Baseline
12 months
24 months
36 months
Difference images between recall samples and baseline sample
Composite resin material wear in an occlusal molar restoration.
■ 100 – 200 µm
■ 200 – 300 µm
■ 300 – 400 µm
□ 400 – 500 µm
Max. depth 490 µm
Volume loss 2.6 mm³

Figure 1
Composite resin material wear in an occlusal molar restoration. Difference images between recall samples and baseline sample measured by means of Laser Scan 3D-Pro, Willytec, Munich, Germany.

surprising, since the phenomenon of wear is very complex and multifactorial (Braem *et al.* 1986). Therefore, long-term quantitative *in vivo* wear measurements are still essential to determine the clinical performance of new posterior composite resin materials (Willems *et al.* 1993a).

3. Wear Rates

Wear rates of dental composite resins should be in the range of *in vivo* enamel wear. Söderholm *et al.* (2001) found that the *in vivo* attritional enamel wear rate in molars was about 39 µm month^{-1} and that the average wear rate on contact-free surfaces was about 9.2 µm month^{-1} with the microscopic measurement technique and 8.5 µm month^{-1} with the laser scanner over a 36-month period.

The wear performance of modern composites is comparable to amalgam and enamel with abrasion wear rates from ~5 µm to ~100 µm per year (Lambrechts *et al.* 1997, Wendt and Leinfelder 1992). Results differ among evaluators because of operator variations, patient variations, and last but not least important, the wear evaluation method (Söderholm *et al.* 1992). The relatively high standard deviations in the results are inherent to *in vivo* studies (Lambrechts *et al.* 1989). The *in vivo* results reveal that the loss of substance is consistently greater in the OCAs than in the CFCAs. Willems *et al.* (1993b) evaluated the wear of five posterior composites at the OCAs and CFOAs in Class II cavities over a 3-yr period with an accurate 3D-measuring technique. A clinical evaluation was also performed. The ultra-fine compact-filled composites showed acceptable OCA-wear rates ranging from 110 μm to 149 μm after 3 yr. This is very similar to the OCA-wear rate of human enamel on molars, which is about 122 μm after 3 yr. The fine compact-filled composite had an unacceptable OCA-wear value of 242 μm after 3 yr. The *ultrafine midway-filled composite* showed an exceptionally high CFOA-wear rate of 151 μm after 3 yr, which gave the impression of it being gradually washed out of the cavity.

The nonlinear wear behavior has been previously discussed by Leinfelder (1988) and may be a result of reduced occlusal stresses as the surface of the composite wears down from the cavosurface margin and becomes somewhat protected by the cavity walls.

For material selection it is only relevant how much time it takes a material to wear to a predefined maximum height loss in comparison to other materials. It has been proposed to convert nonlinear (clinical) height loss data to a linear parameter, introduced as wear life, which is defined as the time it would take a material in a standard restoration to reach a maximum acceptable amount of height loss (Pallav 1996). This wear life should be comparable to the wear life of human enamel.

4. New Technology

In 1998, the first restorative material based on the ormocer technology (Definite [Degussa Agn, Hanau, Germany]) was marketed. Multifunction urethane and tioether(meth)acrylate alkoxysilanes as sol-gel precursors have been developed for synthesis of inorganic–organic copolymer ormocer composites as the dental restorative materials. Ormocers (acronym for organically modified ceramics) are characterized by novel inorganic–organic copolymers in the formulation that allow the modification of mechanical parameters in a wide range (Hickel *et al.* 1998, Manhart *et al.* 2000). According to Xu *et al.* (1999), contact damage resistance of resin composites can be substantially improved by reinforcement by ceramic whiskers fused with silica glass particles. Rudell *et al.* (2002) investigated the use of prepolymerized fused-fiber filler modified composite particles (PP-FFMC). They found that these PP-FFMC particles were better incorporated and resistant to plucking-out of the surface. A current trend in composite technology is toward the use of homogeneously distributed nanoparticles with approved resin matrix protection and reduced surface friction.

Other new filler technology is the use of irregular, spherical or porous filler particles (Luo *et al.* 1998). The significance of these technological developments on *in vivo* wear resistance is still under investigation, but an alternative enamel substitute is thought to be very near at the beginning of the twenty-first century.

5. Conclusion

The ultrafine compact-filled composite resins are very promising materials for posterior use. They have a Young's modulus of elasticity that is higher than that of dentin and have a high content of inorganic particles. They also display good Vickers hardness values compared to dentin and relatively high compressive strength values; these suggest that they are able to support occlusal stresses. Furthermore, their surface roughness ranges from 0.48 μm to 0.71 μm, values very similar to that of enamel. The ultrafine compact-filled composite resins are the materials of choice for restoring posterior cavities at the beginning of the twenty-first century, offering stress-bearing restorations.

The microfilled composite that possesses the highest percent of inorganic filler, the best homogeneous dispersion of filler, and the strongest chemical bond between filler and resin matrix should have the best wear resistance, but not necessary the best fatigue resistance (Lim *et al.* 2002, Willems *et al.* 1993a). A posterior composite resin should also have a good "packability," clinical handling, and possibilities of repair.

Bibliography

Bates J F, Stafford G D, Harrison A 1975 Masticatory function—a review of the literature. 1. The form of the masticatory cycle. *J. Oral Rehabil.* **2**, 281–301

Bayne S C, Taylor D F, Heymann H O 1992 Protection hypothesis for composite wear. *Dent. Mater.* **8**, 305–9

Bayne S C, Taylor D F, Roberson T M, Wilder A D, Sturdevant J R, Heymann H O, Lisk M W 1989 Long term clinical failures in posterior composites. *J. Dent. Res.* **68**, 185 Abstract no. 32

Beatty M W, Swartz M L, Moore B K, Phillips R W, Roberts T A 1993 Effect of cross-linking agent content, monomer functionality, and repeat unit chemistry on properties of unfilled resins. *J. Biomed. Mater. Res.* **27**, 403–13

Beatty M W, Swartz M L, Moore B K, Philips R W, Roberts T A 1998 Effect of microfiller fraction and silane treatment on resin composite properties. *J. Biomed. Mater. Res.* **40**, 12–23

Braem M, Lambrechts P, Van Doren V E, Vanherle G 1986 The impact of composite structure on its elastic response. *J. Dent. Res.* **65**, 648–53

Calais J G, Söderholm K-J M 1998 Influence of filler type and water exposure on flexural strength of experimental composite resin. *J. Dent. Res.* **67**, 836–40

Collins C J, Bryant R W, Hodge K-L V 1998 A clinical evaluation of posterior composite resin restorations: 8-year findings. *J. Dent.* **26**, 311–7

Condon J R, Ferracane J L 1996 Evaluation of composite wear with a new multi-mode oral wear simulator. *Dent. Mater.* **12**, 218–26

Condon J R, Ferracane J L 1997 *In vitro* wear of composite with varied cure, filler level, and filler treatment. *J. Dent. Res.* **76**, 1405–11

Cunningham J, Mair L H, Foster M A, Ireland R S 1990 Clinical evaluation of three posterior composite and two amalgam restorative materials: 3-year results. *Br. Dent. J.* **169**, 319–23

Cvar J F, Ryge G 1971 *Criteria for the Clinical Evaluation of Dental Restorative Materials.* US Government Printing Office, San Francisco USPHS Publication no. 790–244

De Gee A J, Pallav P 1994 Occlusal wear simulation with the ACTA wear machine. *J. Dent.* **22** (suppl. 1), S21–S27

De Gee A J, Pallav P, Davidson C L 1986 Effect of abrasion medium on wear of stress-bearing composites and amalgam *in vitro*. *J. Dent. Res.* **19**, 39–45

Degrange M, Roulet J F 1997 *Minimally Invasive Restorations with Bonding*. Quintessence Publishing, Chicago

Ferracane J L, Condon J R 1992 Post-cure heat treatments for composites: properties and fractography. *Dent. Mater.* **8**, 290–5

Ferracane J L, Condon J R 1999 *In vitro* evaluation of the marginal degradation of dental composites under simulated occlusal loading. *Dent. Mater.* **15**, 262–7

Ferracane J L, Mitchem J C, Condon J R, Todd R 1997 Wear and marginal breakdown of composites with various degrees of cure. *J. Dent. Res.* **76**, 1508–16

Hickel R, Dasch W, Janda R, Tyas M, Anusavice K 1998 New direct restorative materials. *Int. Dent. J.* **48**, 3–16

Hilton T J 2000 Direct posterior esthetic restorations. In: Summitt J B, Robbins J W, Schwartz R S (eds.) *Fundamentals of Operative Dentistry: A Contemporary Approach*, 2nd edn. Quintessence Publishing, Chicago

Jørgensen K D, Asmussen E 1978 Occlusal abrasion of a composite restorative resin with ultra-fine filler—an initial study. *Quintessence Int.* **9**, 73–8

Jørgensen K D, Hørsted P, Janum O, Krogh J, Schultz J 1979 Abrasion of class 1 restorative resins. *Scand. J. Dent. Res.* **87**, 140–5

Juvinall R C, Marshek K M 1991 *Fundamentals of Machine Component Design*, 2nd edn. Wiley, New York

Krejci I, Krejci D, Lutz F 1990 P-30: *In vivo* Untersuchung eines Seitenzahnkomposits während 2,5 Jahren. *Deutsch Zahnärztl Z.* **45**, 773–8

Krejci I, Stergiou G, Lutz F 1991 Einfluss der Nachvergütung auf die Verschleissfestifkeit von Kompositmaterialien. *Deutsch Zahnärztl Z.* **46**, 400–6

Lambrechts P, Braem M, Vanherle G 1985 Accomplishments and expectations with posterior composite resins. In: Vanherle G, Smith D C (eds.) *Proc. Int. Symp. Posterior Composite Resin Dental Restorative Materials*. Peter Szulc Press, Utrecht, The Netherlands, pp. 521–8

Lambrechts P, Braem M, Vanherle G 1987 Evaluation of clinical performance for posterior composite resins and dentin adhesives. *Oper. Dent.* **12**, 53–78

Lambrechts P, Braem M, Vanherle G 1988 Roughness of human enamel as additional acceptance standard for posterior composites. *J. Dent. Res.* **67**, 262 Abstract no. 1192

Lambrechts P, Braem M, Vuylsteke-Wauters M, Vanherle G 1989 Quantitative *in vivo* wear of human enamel. *J. Dent. Res.* **68**, 1752–4

Lambrechts P, Braem M, Vuylsteke-Wauters M, Vanherle G 1991 The surface roughness of enamel-to-enamel contact areas compared with the intrinsic roughness of dental resin composites. *J. Dent. Res.* **70**, 1299–305

Lambrechts P, Vanherle G 1983 Structural evidences of the microfilled composites. *J. Biomed. Mater. Res.* **17**, 249–60

Lambechts P, Vanherle G, Vuylsteke M, Davidson C L 1984 Quantitative evaluation of posterior dental restorations: a new three-dimensional measuring technique. *J. Dent.* **12**, 252–67

Lambrechts P, Willems G, Van Meerbeek B, Perdigão J, Braem M, Vanherle G 1997 Amalgam replacement: can direct resin composite restorations handle the problem? In: Degrange M, Roulet J F (eds.) *Minimally Invasive Restorations with Bonding*. Quintessence Publishing, Chicago, pp. 61–88

Leinfelder K F 1988 Current developments in posterior composite resins. *Adv. Dent. Res.* **2**, 115–21

Lim B S, Ferracane J L, Condon J R, Adey J D 2002 Effect of filler fraction and filler surface treatment on wear of microfilled composites. *Dent. Mater.* **18**, 1–11

Luo J, Lannutti J J, Seghi R R 1998 Effect of filler porosity on the abrasion resistance of nanoporous silica gel/polymer composites. *Dent. Mater.* **14**, 29–36

Lutz F, Philips R W, Roulet J C, Setcos J C 1984 *In vivo* and *in vitro* wear of potential posterior composites. *J. Dent. Res.* **63**, 914–20

Manhart J, Kunzelmann K H, Chen H Y, Hickel R 2000 Mechanical properties and wear behavior of light-cured packable composite resins. *Dent. Mater.* **16**, 33–40

Mazer R B, Leinfelder K F, Russell C M 1992 Degradation of microfilled posterior composite. *Dent. Mater.* **8**, 185–9

McKinney J E, Wu W 1982 Relationship between subsurface damage and wear of dental restorative composites. *J. Dent. Res.* **61**, 1083–8

Mehl A, Gloger W, Kunzelmann K H, Hickel R 1997 A new optical 3-D device for the detection of wear. *J. Dent. Res.* **76**, 1799–807

Mitchem J C, Gronas D G 1990 The use of dentures for the evaluation of the abrasive wear of composites. *J. Dent. Res.* **69**, 126 Abstract

Miyasaka T, Yoshida T 2000 Effect of binary and ternary filler mixtures on the mechanical properties of composite resins. *Dent. Mater.* **19**, 229–44

Mohsen N M, Craig R 1995 Effect of silanization of fillers on their dispersability by monomer system. *J. Oral Rehabil.* **22**, 183–9

Pallav P 1996 Occlusal wear in dentistry, fundamental mechanisms, clinical implications, and laboratory assessment. Ph.D. thesis, Academisch Proefschrift, Academisch Cerbrum Tandheelkunde Amsterdam

Pallav P, de Gee A J, Davidson C L, Erickson R L, Glasspoole E A 1989 The influence of admixing microfiller to small-particle composite resin on wear, tensile strength, hardness, and surface roughness. *J. Dent. Res.* **68**, 489–90

Roulet J F 1987 A material scientist's view: assessment of wear and marginal integrity. *Quintessence Int.* **18**, 543–52

Roulet J F 1988 The problems associated with substituting composite resins for amalgam: a status report on posterior composites. *J Dent.* **16**, 101–13

Rudell D E, Maloney M M, Thompson J Y 2002 Effect of novel filler particles on the mechanical and wear properties of dental composites. *Dent. Mater.* **18**, 72–80

Sakaguchi R L, Douglas W H, DeLong R, Pintado M R 1986 The wear of a posterior composite in an artificial mouth: a clinical correlation. *Dent. Mater.* **2**, 235–40

Shirai K, Yoshida Y, Nakayama Y, Fujitani M, Shintani H, Wakasa K, Snauwaert J, Van Meerbeek B 2000 Assessment of decontamination methods as pretreatment of silanization of composite glass fillers. *J. Biomed. Mater. Res.* **53**, 204–10

Söderholm K-J M 1985 Filler systems and resin interface. In: Vanherle G, Smith D C (eds.) *Proc. Int. Symp. Posterior Composite Resin Dental Restorative Materials*. Peter Szulc Press, Utrecht, The Netherlands, pp. 139–60

Söderholm K-J M, Lambrechts P, Sarett D, Abe Y, Yang M C K, Labella R, Yildiz E, Willems G 2001 Clinical wear performance of eight experimental dental composites over three years determined by two measuring methods. *Eur. J. Oral Sci.* **109**, 273–81

Söderholm K-J M, Richard N D 1998 Wear resistance of composites: a solved problem? *General Dentistry* **46**, 256–63

Söderholm K-J M, Roberts M J, Antonson D E, Anusavice K J, Mauderli A P, Sarrett D C, Warren J W 1992 Visual and profilometric wear measurements. *Acta Odontol. Scand.* **50**, 121–7

Söderholm K-J M, Shang S W 1993 Molecular orientation of silane at the surface of colloidal silica. *J. Dent. Res.* **72**, 1050–4

Suzuki S, Leinfelder K F, Kamai K, Tsuchitani Y 1995 Effect of particle variation on wear rates of posterior composites. *Am. J. Dent.* **8**, 173–8

Taylor D F, Bayne S C, Leinfelder K F, Davis S, Koch G C 1994 Pooling of long term clinical wear data for posteror composites. *Am. J. Dent.* **7**, 167–74

Truong V T, Tyas M J 1988 Prediction of *in vivo* wear in posterior composite resins: a fracture mechanics approach. *Dent. Mater.* **4**, 318–27

Venhoven B M A, de Gee A J, Werner A, Davidson C L 1996 Influence of filler parameters on the mechanical coherence of dental restorative resin composites. *Biomaterials* **17**, 735–40

Wang J-C, Charbeneau G T, Gregory W A, Dennison J B 1989 Quantitative evaluation of approximal contacts in Class 2 composite resin restorations. *Oper. Dent.* **14**, 193–202

Wendt S L, Leinfelder K F 1992 Clinical evaluation of Clearfil photoposterior: 3-year results. *Am. J. Dent.* **5**, 121–5

Willems G, Lambrechts P, Braem M, Vanherle G 1993a Composite resins in the 21st century. *Quintessence Int.* **24**, 641–58

Willems G, Lambrechts P, Braem M, Vanherle G 1993b Three-year follow-up of five posterior composites: *in vivo* wear. *J. Dent.* **21**, 74–8

Willems G, Lambrechts P, Lesaffre E, Braem M, Vanherle G 1993c Three-year follow-up of five posterior composites: SEM study of differential wear. *J. Dent.* **21**, 79–86

Xu H H K, Martin T A, Antonucci J M, Eichmiller F C 1999 Ceramic whisker reinforcement of dental resin composites. *J. Dent. Res.* **78**, 706–12

Yap A U, Chew C L, Ong L F, Teoh S H 2002 Environmental damage and occlusal contact area wear of composite restoratives. *J. Oral Rehabil.* **29**, 87–97

Yoshida Y, Shirai K, Nakayama Y, Itoh M, Okazaki M, Shintani H, Inoue S, Lambrechts P, Vanherle G, Van Meerbeek B 2002 Improved filler-matrix coupling in resin composites. *J. Dent. Res.* **81**, 270–3

K. Goovaerts, P. Lambrechts, J. De Munck,
L. Bergmans and B. Van Meerbeek
Catholic University of Leuven, Belgium

Composite Materials, Microstructural Design of

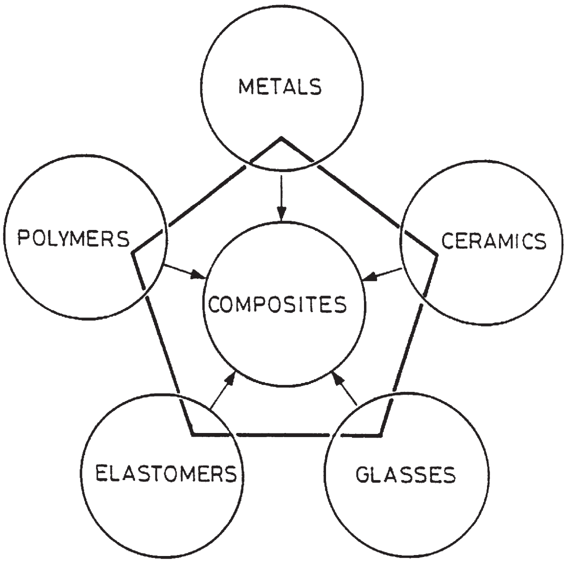


Figure 1
The material classes from which composites are made.

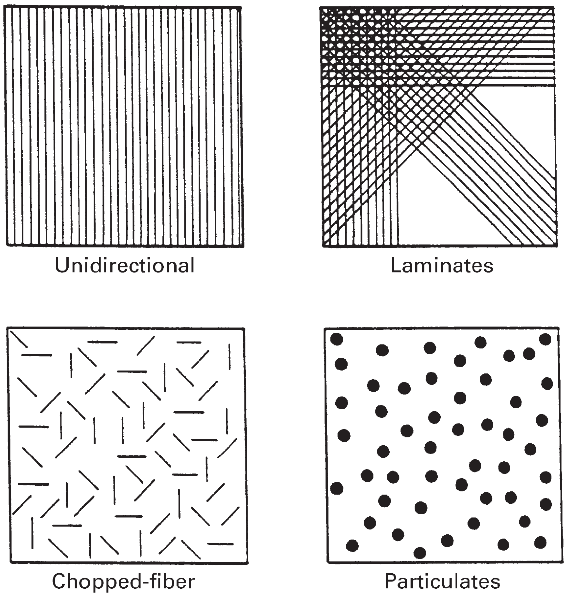


Figure 2
Unidirectional fibrous composites, laminated composites, chopped-fiber and particulate composites.

Any two materials could, in principle, be combined to make a composite (Fig. 1), and they might be mixed in many geometries (Fig. 2). This article summarizes a scheme for identifying the component materials that could be used to make composites with potentially attractive properties. Three concepts are used. The first is that of performance indices, which isolate the combination of material properties that maximize performance; the second is that of materials-selection charts, onto which both material properties and performance indices can be plotted; and the third is the use of bounds to define the envelope of properties accessible to a given composite system (Ashby 1993).

1. Indices, Selection Charts, and Property Bounds

A performance index is a property or group of properties that measures the effectiveness of a material in performing a given function (Ashby 1999). The best material for a light, stiff tie (a tensile member) is that with the greatest specific stiffness, E/ρ, but it is not always so straightforward. The lightest beam (a member loaded in bending) with a prescribed stiffness is that made of the material with the greatest value of $E^{1/2}/\rho$; the best springs are those made of materials with the highest values of σ_f^2/E, where σ_f is the yield or fracture stress; the precision device that is least distorted by heat is that made of the material with the highest value of λ/α, where λ is the thermal conductivity and α is the expansion coefficient. The performance indices are $E^{1/2}/\rho$, σ_f^2/E and λ/α. The best choice of material is that with the largest value of the appropriate index. The indices provide a set of pointers; they direct the composite-developer towards material combinations that, potentially, offer something new.

The indices, frequently, combine two or more properties. This suggests the idea of constructing charts on which one property is plotted against another in such a way that both the properties and their combinations can be examined. Figure 3 illustrates this. The axes are Young's modulus E and density ρ. The scales are logarithmic, and span a range so wide that almost all engineering materials are included. The index E/ρ appears on these axes as a family of parallel lines of slope 1; the index $E^{1/2}/\rho$ appears as a family of slope 2. Other charts give optimal selection via other indices. Charts and indices have been fully described by Ashby (1999).

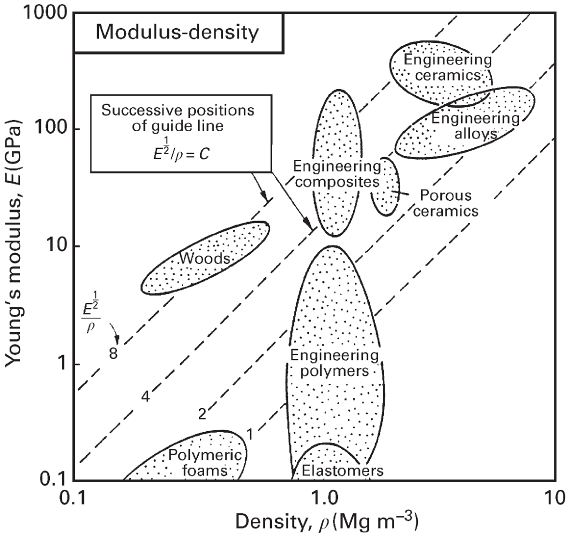


Figure 3
A schematic selection chart: modulus E is plotted against density ρ.

On a macroscopic scale a composite behaves like a homogenous solid with its own set of thermomechanical properties. Calculating these precisely is difficult. It is much easier to bracket them by bounds or limits: upper and lower values between which the properties lie. For most properties, this can be done accurately enough to identify interesting possibilities.

Seven mechanical and thermal properties are of primary interest in assessing the potential of a new composite: density, modulus, strength, toughness, thermal conductivity, expansion coefficient, and heat capacity (Table 1); others, like fracture toughness and thermal diffusivity, are calculated from them. In this section we assemble bounds or limits for material properties. The term "bound" will be used to describe a rigorous boundary, one that the value of the property *cannot*—subject to certain assumptions—exceed or fall below. However, it is not always possible to derive bounds; then the best that can be done is to derive "limits" outside which it is *unlikely* that the value of the property will lie. In what follows, density and specific heat are calculated exactly; moduli, thermal expansion, thermal conductivity and diffusivity are bracketed by bounds, but strength and toughness can only be enclosed by limits. As far as possible, all are based on micromechanical modeling; their origins have been detailed by Hull (1981), Kelly and Macmillan (1986), Chamis (1987), and Ashby (1993). The important point is that the bounds or limits bracket the properties of all arrangements of matrix and reinforcement shown in Fig. 2; by using them we escape from the need to model individual geometries.

1.1 Density

When a volume fraction f of a reinforcement r (density ρ_r) is mixed with a volume fraction $(1-f)$ of a matrix m (density ρ_m) to form a composite with no residual porosity or void-space, the composite density

Table 1
Design limiting properties.

Property	Symbol and usual units
Density	ρ (Mg m^{-3})
Young's modulus	E (GPA)
Strength	σ_f (MPa, sometimes GPa)
Toughness; or fracture toughness	J_{1c} (kJ m^{-2}); K_{1c} (MPa m$^{1/2}$)
Linear thermal expansion coefficient	α (K^{-1})
Specific heat	C_p (J kg^{-1} K^{-1})
Thermal conductivity	λ (W m^{-1} K^{-1})
Thermal diffusivity	a (m^2 s^{-1})

is given exactly by a rule of mixtures (an arithmetic mean, weighted by volume fraction)

$$\rho_r = f\rho_r + (1-f)\rho_m \tag{1}$$

1.2 Modulus

The modulus of a composite is bracketed by the well-known Voigt and Reuss bounds:

$$E_r = fE_r + (1-f)E_m \tag{2}$$

and

$$E_r = \frac{E_m E_r}{fE_m + (1+f)E_r} \tag{3}$$

Here E_r is the Young's modulus of the reinforcement and E_m that of the matrix. Tighter bounds exist, but Eqns. (2) and (3) are sufficient for our purpose.

1.3 Strength

The upper bound is, as with modulus, a rule of mixtures

$$(\sigma_f)_u = f(\sigma_f)_r + (1-f)(\sigma_f)_m \tag{4}$$

where $(\sigma_f)_m$ is the strength of the matrix and $(\sigma_f)_r$ is that of the reinforcement. A lower limit is the yield strength of the matrix enhanced slightly by the plastic constraint imposed by the reinforcement

$$(\sigma_f)_1 = (\sigma_f)_m 1 + \frac{1}{16}\left(\frac{f^{1/2}}{1-f^{1/2}}\right) \tag{5}$$

The bounds are wide, but they still allow important conclusions to be reached.

1.4 Specific Heat

The specific heats of solids at constant pressure, C_p, are almost the same as those at constant volume, C_v. If they were identical, the heat capacity per unit volume of a composite would, like the density, be given exactly by a rule-of-mixtures

$$\rho C_p = f\rho_r(C_p)_r + (1-f)\rho_m(C_p)_m \tag{6}$$

where $(C_p)_r$ is the specific heat of the reinforcement and $(C_p)_m$ is that of the matrix.

1.5 Thermal Expansion Coefficient

We use the approximate lower bounds of Levin

$$\alpha_L = \frac{E_r\alpha f + E_m\alpha_m(1-f)}{E_r f + E_m(1-f)} \tag{7}$$

and the upper bound of Schapery

$$\begin{aligned}\alpha_u = {} & f\alpha_r(1+v_r) + (1-f)\alpha_m(1+v_m) \\ & - \alpha_L[fv_r + (1-f)v_m]\end{aligned} \tag{8}$$

where α_r and α_m are the two expansion coefficients and v_r and v_m the Poisson's ratios.

1.6 Thermal Conductivity

A composite containing parallel continuous fibers has a conductivity, parallel to the fibers, given by a rule-of-mixtures

$$\lambda_u = f\lambda_r + (1-f)\lambda_m \tag{9}$$

This is an upper bound; in any other direction the conductivity is lower. The transverse conductivity of a parallel-fiber composite (again assuming good bonding and thermal contact) lies near the lower bound first derived by Maxwell

$$\lambda_L = \left(\frac{\lambda_r + 2\lambda_m - 2f(\lambda_m - \lambda_r)}{\lambda_r + 2\lambda_m + f(\lambda_m - \lambda_r)}\right) \tag{10}$$

Particulate composites, too, have a conductivity near this bound. Details are given by Ashby (1993), in which bounds and limits for thermal diffusivity, toughness, and fracture toughness are also discussed.

2. Applications: The Potential of Composite Systems

The last step is that of combining the indices, the charts and the bounds to design composites for specific applications. Two examples are developed, but it will be clear that the method is a general one, and can be extended further.

2.1 Composite Design for Stiffness at Minimum Weight

Consider, first, design of a composite for a light, stiff, beam of fixed section-shape, to be loaded in bending. The efficiency is measured by the index $M = E^{1/2}/\rho$. Imagine, as an example, that the beam is at present made of an aluminum alloy and that the alloy could be stiffened by incorporating particles or fibers of beryllium (Be) or of alumina (Al_2O_3) in it. Both are much stiffer than aluminum—that is, they have higher moduli.

Figure 4 is a small part of the E–ρ chart introduced in the last section. Three groups of materials are shown: aluminum and its alloys; beryllium; and a range of aluminas. Composites made from them have densities given exactly by Eqn. (1), and moduli that are bracketed by the bounds of Eqns. (2) and (3). Both of these moduli depend on volume fraction of

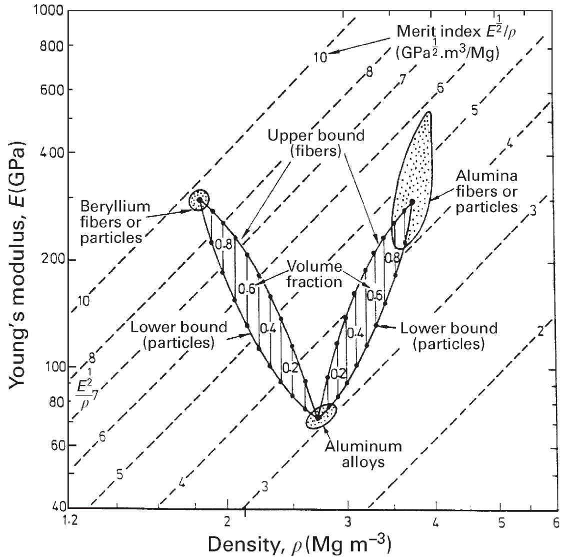


Figure 4
Part of a modulus–density space, showing aluminum alloys, beryllium and alumina. The moduli of the Al–Be and the Al–Al_2O_3 composites are bracketed by the bounds of Eqns. (1), (2), and (3). The Al–Be composite offers much greater gains in M than the Al_2O_3 composite.

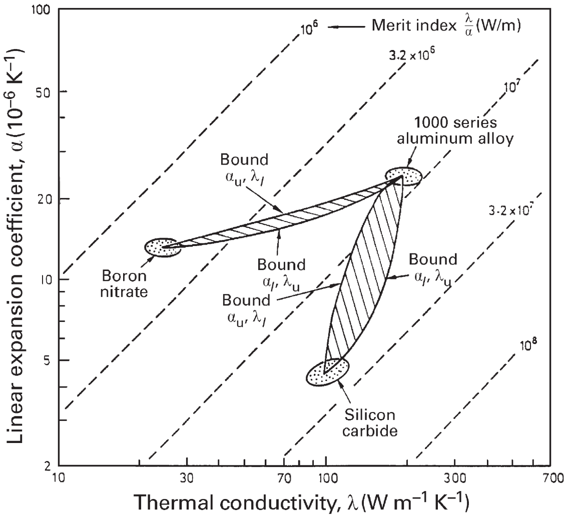


Figure 5
A part of expansion coefficient–thermal conductivity space, showing aluminum alloys, boron nitride and silicon carbide. The properties of Al–BN and Al–SiC composites are bracketed by the bounds described by Eqns. (7)–(10). The Al–SiC composites have better values of the index M (which increases towards the bottom right) than the Al–BN composites.

reinforcement, and through this, on density. Upper and lower bounds for the modulus–density relationship can thus be plotted on the $E-\rho$ chart using volume fraction f as a parameter, as shown in Fig. 4. Any composite made by combining aluminum with beryllium will have a modulus that lies somewhere in the Al–Be envelope; any made of aluminum and alumina will have a modulus contained in the envelope for Al–Al_2O_3. Fibrous reinforcement gives a longitudinal modulus (that parallel to the fibers) near the upper bound; particulate reinforcement or transversely loaded fibers give moduli near the lower one.

Superimposed on Fig. 4 is a grid showing the performance indices M—the quantity we wish to maximize. The bound-envelope for Al–Be composites extends almost normal to the grid, while that for Al–Al_2O_3 is, initially, parallel to the M grid: 30% of particulate Al_2O_3 gives almost no gain in M. The underlying reason is clear: both beryllium and Al_2O_3 increase the modulus, but only beryllium decreases the density; both indices are more sensitive to density than to modulus.

2.2 Composite Design for Specific Thermal Properties

Thermo-mechanical design involves the specific heat, C_p, the thermal expansion, α, the conductivity, λ, and the thermal diffusivity, $\alpha = \lambda(\rho C_p)^{-1}$. These composite properties are bounded by Eqns. (6)–(10). They are involved in a number of performance indices. One is the criterion for minimizing thermal distortion $M-\lambda/\alpha$.

Figure 5 shows a small part of the $\lambda-\alpha$ materials selection chart, with a grid of lines of the index M superimposed on it. Three groups of materials are shown, aluminum alloys, boron nitride (BN), and silicon carbide (SiC). The thermal properties of Al–BN and Al–SiC are bracketed by lines that show the bounding equations. The plot reveals immediately that SiC reinforcement in aluminum increases performance (as measured by M); reinforcement with BN decreases it. By systematically testing alternative combinations of material and reinforcement in this way, the pair that offers the greatest potential gain in performance can be identified.

3. Summary and Conclusions

Developing a new composite is a long and expensive business. It is helpful, before starting, to have an idea of what its strengths and weaknesses might be and where its applications might lie. This article outlines a procedure for doing this. It combines the ideas of performance indices, materials selection charts and bounds or limits for composite properties to identify the composites which, potentially, have attractive combinations of mechanical and thermal properties. Two examples of the method are given in the text. The same approach can be applied to a wide range of thermal and mechanical applications.

See also: Fibrous Reinforcements for Composites: Overview*; Polymer Matrix Composites, Recycling of*; Composites, Physical Properties of*; Continuous Parallel Fiber Composites: Deformation and Strength*; Polymer Matrix Composites with Roughly Equiaxed Reinforcements: Microstructure and Mechanical Behavior*; Ceramic Matrix Composites with Roughly Equiaxed Reinforcements: Microstructure and Mechanical Behavior*; Metal Matrix Composites with Roughly Equiaxed Reinforcements: Microstructure and Mechanical Behavior*; Laminates: Physical and Mechanical Behavior*; Designing with Composites*; Polymer Matrix Composites: Applications*; Metal Matrix Composites: Applications*; Ceramic Matrix Composites: Applications*

Bibliography

Ashby M F 1993 Criteria for selecting the components of composites. *Acta Metall. Mater*. **41**, 1313–35

Ashby M F 1999 *Materials Selection in Mechanical Design*, 2nd edn. Butterworth Heinemann, Oxford, UK

Chamis C C 1987 In: Weeton J W, Peters D M, Thomas K L (eds.) *Engineers Guide to Composite Materials.* American Society for Metals, Metals Park, OH, pp. 3-8–3-24

Hull D 1981 *An Introduction to Composite Materials*. Cambridge University Press, Cambridge, UK

Kelly A, Macmillan 1986 *Strong Solids*, 3rd edn. Oxford University Press, Oxford, UK

M. F. Ashby
University of Cambridge, UK

Composite Materials: Environmental Effects

Composite materials must, as with all materials, resist chemical attack by their service environment. The characteristics of the reinforcements and of the matrix materials can be altered by the absorption of agents, by various mechanisms of corrosive attack, or by the temperatures at which the materials are used.

With composite materials these problems can be amplified because their behavior depends not only on the characteristics of their components, which are the reinforcements and the matrix materials, but also, and often most critically, on the quality of the interfaces between the components. It is the interface that determines load transfer in resin and metal matrix composites, and its controlled failure is a major factor in determining the toughness of ceramic matrix composites. Often the interfaces are vulnerable to outside effects and provide preferential channels for the ingress of agents which can modify the interactions between the fibers and the matrix. In this way composites sometimes may be dramatically deteriorated and their properties reduced by the environment in which they are used, even though both matrix and reinforcement offer good environmental resistance.

In the case of resin matrix composites, the most serious concern for their environmental aging is the effect of water, as the presence of water molecules at the fiber–matrix interface can lead to the failure of the chemical bonds assuring the adhesion of the fibers to the matrix. Such a failure is irreversible and strikes at the basic mechanism of fiber reinforcement since it removes the ability of the matrix to transfer load to the fibers. This is made worse as the water at the interface acts as a lubricant. Furthermore, water is everywhere: applications for which the composite must be immersed in water obviously raise the question of its effects, but water is also in the air in the form of water vapor. Resin matrix composites will absorb water to a lesser or greater extent and although measures such as painting can hinder its penetration, eventually the composite will either come to an equilibrium with its environment or will be degraded by it.

It is therefore necessary to understand and to be able to calculate the rate of diffusion of water in these materials as a function of water concentration and temperature, as the process is thermoactivated. Water also has the effect of a lubricant at the molecular level on the bulk structure of the resin which is revealed most strikingly by a fall in glass transition temperature, T_g, as absorption proceeds. In some resin systems this can lead to dramatic falls in T_g so that at the beginning of exposure the composite may be well below T_g but as water is absorbed T_g falls below the ambient temperature and as a consequence the physical behavior of the matrix is greatly changed. This effect can be totally or partially reversible. Water can also provoke the hydrolysis of the resin leading to a deterioration of the matrix and hence of the composite. In the case of thermoplastic matrix composites, this can lead to an increase in crystallinity and progressive embrittlement of the matrix. In some high-temperature resin systems, oxidation of the polymer matrix can also be a source of degradation.

In the case of metal matrix composites it is often the interaction between the fibers and matrix during composite manufacture which can be the cause of deterioration of the material by corrosion. Also, when their reinforcement can conduct electricity (notably carbon fibers), metal matrix composites can be subject to rapid corrosive attack of one of their phases when water is present and the two types of materials possess markedly different electron energy levels.

Ceramic matrix composites rely on the interface to act as a mechanical fuse which fails when a crack approaches and so blunts further propagation. Even when the crack passes around the fiber the failure of the interface over a substantial distance provides an important mechanism for absorbing energy and so contributes to the overall toughness of the composite. The interface in such composites is controlled so as to facilitate this debonding process. Their use at high temperatures in air, however, can cause major changes to the interfaces, which make the interfaces strong and, in turn, the composite brittle so that any crack provokes fiber failure and interfacial debonding does not occur. In this case, the increased toughness of the ceramic matrix composite with respect to the monolithic ceramic is lost. In addition, the fibers themselves evolve at high temperatures so that they may lose strength and begin to creep.

1. Water Diffusion in Polymer Composites

Diffusion laws have generally been applied to the modeling of simple water uptake, and it has been shown that Fick's law can be applied to carbon-fiber-reinforced epoxy resin subjected to humid environments (Shen and Springer 1976). Water is considered to remain in a single free phase driven to penetrate the resin by the water concentration gradient. Other studies have indicated, however, that non-Fickian processes do occur which can complicate the understanding of the role of the water and which may lead to irreversible changes in mechanical properties (Dewimille and Bunsell 1982). In particular, water ingress can occur by capillary action along the fiber–matrix interface, and some resin systems can be dissolved by water.

Water penetration can cause swelling and plasticization of the resin. Composite structures consisting of several layers of differing types of fiber layup, often with a thick resin or gel coat, can suffer from water being trapped in a particular layer or at the interface between layers. This effect may be irreversible because of osmosis, and blistering may result (Marine *et al.* 1988). This is a particular problem with glass-fiber-reinforced boats for which considerable efforts are made, by a judicious choice of resin materials and manufacturing techniques, to avoid this type of damage. In addition to changes in mechanical properties, modification of other physical characteristics may occur, such as changes in dielectric properties.

1.1 Diffusion Models

(*a*) *Fick's law and single-phase diffusion.* Fick's law assumes that the water remains in a single free phase which is driven to penetrate the material by the water concentration gradient (Crank 1956), $\mathrm{d}c/\mathrm{d}x$, so that the water uptake as a function of time is given by:

$$\frac{\partial c}{\partial t} = \frac{\mathrm{d}}{\mathrm{d}x}\left(D\frac{\mathrm{d}c}{\mathrm{d}x}\right) \tag{1}$$

where D is a diffusion coefficient considered to be constant in the direction of diffusion, x, so that Eqn. (1) can be rewritten as:

$$\frac{\partial c}{\partial t} = D_x\frac{\mathrm{d}^2c}{\mathrm{d}x^2} \tag{2}$$

It has been shown (Crank 1956) that for an infinitely large plate of thickness h the mass of water, M_t, absorbed in time t across a unit surface area as a percentage of the mass of water absorbed at saturation, M_∞, is given by:

$$\frac{M_t}{M_\infty} = 1 - \frac{8}{\pi^2}\sum_{n=0}^{\infty}\frac{1}{(2n+1)^2}\exp\left[-D(2n+1)^2\pi^2t/h^2\right] \tag{3}$$

This equation has been simplified by Shen and Springer (1976) to show that the initial absorption is given by:

$$\frac{M_t}{M_\infty} = \frac{4}{h}\left(\frac{Dt}{\pi}\right)^{1/2} \tag{4}$$

The initial linear relationship of M_t/M_∞ as a function of the square root of time is to be noted, and so water uptake is usually plotted as M_t/M_∞ vs. square root of time. Figure 1 shows the classical Fickian behavior of a glass-fiber-reinforced epoxy resin at two temperatures and at different values of relative humidity. The composite reaches equilibrium with its environment and saturates. It can be seen that the saturation levels are independent of temperature but do depend on relative humidity. Figure 2 shows that the diffusivity, D, obeys a thermoactivated Arrhenius-type relationship and is independent of relative humidity. This behavior indicates that the water is absorbed by the resin and is linked to it by hydrogen bonds.

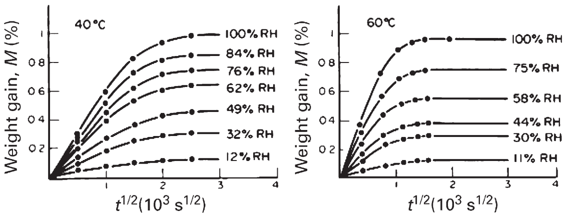


Figure 1
Absorption curves obtained under different conditions of temperature and relative humidity with glass-fiber-reinforced DGEBA epoxy resin, showing classical Fickian absorption.

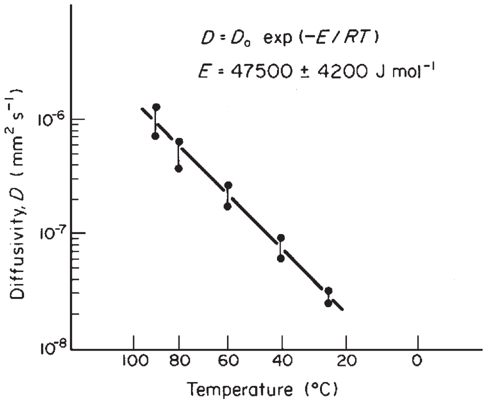


Figure 2
Thermoactivated diffusion into a glass-fiber-reinforced DGEBA epoxy resin, obeying an Arrhenius relationship with temperature. E is the activation energy.

Under most conditions water can penetrate into the material in several directions. In general, the coefficient of diffusivity varies with the direction so that edge effects can become important and should be considered in any analysis of the water uptake of structures which present more than one surface to water diffusion.

For a unidirectional specimen, two diffusion coefficients must be considered accounting for penetration normal ($D_\perp$) and parallel ($D_\parallel$) to the fibers. A third coefficient (D_r) for the unreinforced resin also exists.

(*b*) *The Langmuir model of diffusion.* In some cases the simple Fickian model proves inadequate, such as when two plateaus in the absorption curve of mass of water absorbed vs. square root of time are observed or where it is found that the absorption depends on the specimen thickness. It is then sometimes useful to consider a more complex model of the Langmuir type in which the water is considered to be in two phases, one free to diffuse and the other trapped in the absorbing medium. This model has been treated by Crank (1956). In addition to the two parameters D and M_∞ used in the simple free-phase diffusion model, the two-phase model introduces two other parameters: α, the probability of a trapped water molecule being released and γ, the probability of a free water molecule being trapped.

The analysis of this situation results in the following simplified equation which describes the rate of water uptake from both sides in a plate of thickness $2l$:

$$\frac{M_t}{M_\infty} = 1 - \frac{\gamma}{\gamma+\alpha}e^{-\alpha t} - \frac{8}{\pi^2}\frac{\alpha}{\gamma+\alpha}e^{-\gamma t}\sum_{n=1}^{\infty}\frac{1}{(2n-1)}e^{-(\pi/2)^2 Dt} \tag{5}$$

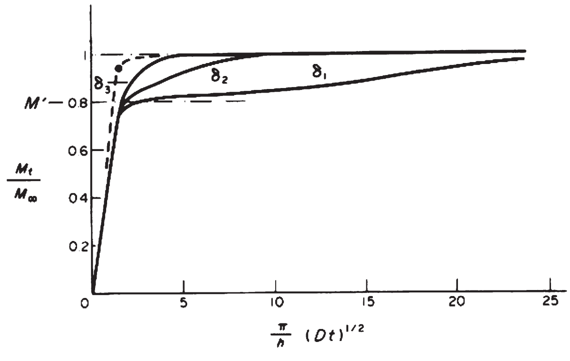


Figure 3
Two-phase diffusion as a function of square root of time, as described by the Langmuir model, showing the influence of specimen thickness. $\delta = (h/\pi)(\alpha/D)^{1/2}$, $\delta_1 = 0.05$, $\delta_2 = 0.2$, and $\delta_3 = 0.5$, where h is the specimen thickness. The dashed curve follows Fick's law ($\alpha/(\alpha + \gamma) = M'/M_x = 0.8 \rightarrow \alpha/\gamma = 4$).

Figure 3 shows how Eqn. (5) describes the uptake of water into plates of differing thicknesses, h. It should be noted that the Langmuir model reduces to the simple Fickian model when $\gamma = 0$.

In passing to a two-phase model it is assumed that the absorbed water is in two states. The free phase having a concentration $n(t)$ water molecules per unit volume can diffuse in the resin, but can be linked or trapped in the resin with a probability γ, and the combined phase with concentration $N(t)$ able to break free with a probability α. At saturation:

$$n_\infty \gamma = N_\infty \alpha \tag{6}$$

i.e., there is an equilibrium between those water molecules being trapped and those being freed.

The physical processes involved in water absorption by the composite are certainly more complex than the processes considered in the Langmuir model but because of its additional flexibility the model is applicable to a wider range of cases.

1.2 Physicochemical Aspects of Absorption in Resin Matrix Composites

Most polymers absorb water to a greater or lesser degree so that, even in the absence of defects or preferential routes for water uptake such as are provided by poor fiber–matrix adhesion, resin matrix composites will absorb water in humid environments. Water diffusion occurs in most resins by hydrogen bonds being formed between the water and the molecular structure of the resin.

Epoxy resins are the most widely used resins in high-performance composite structures. They are mixed with a hardener, which can be of several types, to produce a cross-linked structure which provides many sites for hydrogen bonding for the water molecules, such as hydroxyl groups (O–H), phenol groups (O–C), amine groups (N–O), and sulfone groups (O–S). Figure 4 shows the sites available in a commonly used epoxy consisting of a bisphenol-A epoxy resin cross-linked with a dicyandiamide hardener. The hydrolysis of epoxy resins is in itself a cause of degradation and is not by any means limited to this class of matrix materials. Certain fibers, such as those of glass, can be damaged by prolonged exposure to water which arrives at the interface. Glasses are made up of silica in which are dispersed metallic oxides including those of the alkali metals. These latter nonsilicate constituents represent microheterogeneities which are hygroscopic and hydrolyzable. The absorption of water by glass is therefore characterized by the hydration of these oxides. The most common form of glass fiber is E-glass, which contains only small amounts of alkali metal oxides and so is resistant to damage by water.

However, the presence of water at the glass-fiber surface may lower its surface energy and promote crack growth. Water trapped at the interface may also allow components of the resin to go into solution and if an acidic environment is formed the fiber can be degraded. Organic fibers such as the aramid family can absorb considerable quantities of water and although this does not lead to a deterioration of fiber properties, the resulting swelling of the fiber may be a cause of composite degradation. The size, or coating, put onto many fibers to protect them from abrasion and to ensure bonding with the matrix also serves to protect them and the composite from damage from water absorption by eliminating sites at which the water can accumulate. However, this is not always adequate.

(*a*) *Influence of the resin matrix on composite degradation.* Degradation of glass fibers embedded in an epoxy matrix depends greatly on the type of hardener used to cross-link the resin. Figure 1 shows the classical Fickian behavior of a glass-fiber-reinforced bisphenol-A epoxy cross-linked with a diamine hardener. The saturated composite reveals a somewhat increased tenacity owing to a softening of the matrix which is accompanied by a fall in glass transition temperature (Bonniau and Bunsell 1981, Springer 1982). Damage does occur under severe conditions (90 °C, 100% relative humidity for two weeks) and appears as a progressive whitening of the material with no detectable weight gain. This is probably due to microcracking which begins at the surface.

A change to a dicyandiamide hardener produces, however, a significant change in water-uptake kinetics, and this behavior can only be explained by the two-phase model of absorption. Saturation is attained after considerably longer exposure than with the diamine-hardened resin, but after equilibrium is reached both composites behave in a similar manner.

The use of an anhydride as the cross-linking agent, which is extremely common, produces a composite which, in water that is warmer than about 40 °C, degrades quite rapidly. Saturation is not achieved and severe weight loss after drying is observed, owing to leaching of the resin into the water. An excess of the anhydride from the cross-linking process combines with the water to produce an acid which degrades the composite. The correct choice of hardener in these epoxy composites is therefore extremely important if they are to be used in warm, humid conditions.

The glass transition temperature, T_g, of organic matrix composites is lowered on water uptake. In the case of thermosetting resins such as epoxides, this decrease can be as much as 30 °C and is even more in the case of glass-fiber-reinforced polyamide, as shown in Fig. 5.

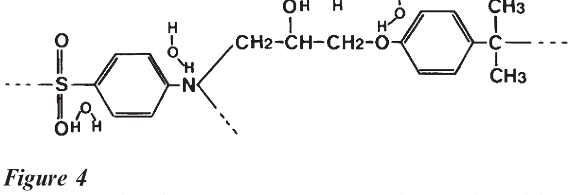


Figure 4
Sites available for water molecules to be combined by hydrogen bonding onto the molecular structure of a DGEBA epoxy resin cross-linked with a dicyandiamide hardener.

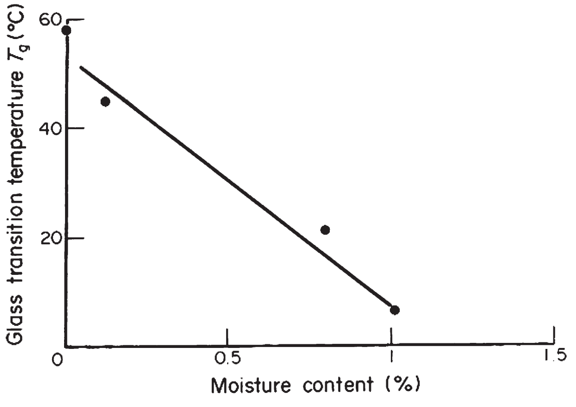


Figure 5
Glass transition temperature of glass-fiber-reinforced nylon decreasing dramatically with water uptake. The composite is immersed in water at 90 °C and T_g is measured using a differential scanning calorimeter.

Families of resins able to withstand higher temperatures than the epoxy resins have been developed and in general are less susceptible to water uptake. Prolonged exposure of these resins to hot water can, however, lead to irreversible changes. Polystyrylpyridine (PSP) reinforced with carbon fibers can operate at 250 °C for 1000 hours. Prolonged exposure to humid conditions, e.g., 70 °C and 100% relative humidity, reveals an apparent initial Fickian absorption leading to saturation. Longer exposures lead, however, to further water uptake and the creation of microcracks as well as the hydrolysis of the polymer matrix. Thermal cycling from 70 °C to 150 °C in the presence of water increases water absorption owing to the water being vaporized and forcing microcracks further open. Cycling to 250 °C leads to a more marked worsening of properties owing to greater damage exacerbated by oxidation of the resin. The effects of water on other high-temperature resins, such as PMR-15 and Avimid-N which are rated to be used above 300 °C when dry, are to produce a fall in T_g of around 100 °C during hygrothermal cycling.

1.3 Mechanical Degradation owing to Water Absorption in Resin Matrix Composites

As mentioned above, water penetration into resin matrix composites can lead to changes in mechanical properties. A softening of the matrix is commonly observed which in some cases can be beneficial, inducing a greater toughness. However, changes are usually associated with deterioration in properties such as strength or modulus. These changes can appear to be reversible as, upon drying, the composite may regain most of its lost properties. Figure 6 illustrates this for a unidirectional glass-reinforced epoxy subjected to three-point bending. It can be seen that drying the specimen almost reversed the effects seen at saturation. A second cycle in humid conditions usually reveals that water penetration occurs much more quickly and property loss is much more rapid than in the first cycle. This behavior is because of the destruction of the chemical bonding at the interface by the presence of water. Upon drying, load transfer between fiber and matrix is assured by frictional forces which may largely compensate for the loss of chemical bonding. A second exposure to humidity results in rapid water uptake as the water can quickly penetrate the composite along the fiber–matrix interfaces and, acting as a lubricant, rapidly results in loss of mechanical properties.

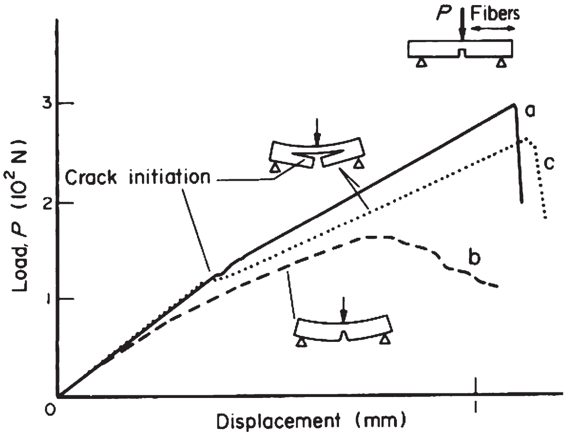


Figure 6
Load–deflection curves obtained with notched three-point bending tests on a unidirectional glass-fiber-reinforced epoxy resin: (a) initial state, (b) after 96 hours in boiling water, and (c) after 96 hours in boiling water followed by drying.

2. Corrosion of Metal Matrix Composites (MMCs)

Fiber- or particle-reinforced metals are made in much smaller volumes than polymer composites but are of interest for applications that require higher operating temperatures. In addition, MMCs do not suffer from outgassing of water vapor or solvents, which cause major problems in space applications, as do resin composites. They can be made with very low coefficients of thermal expansion by balancing the properties of the reinforcements and those of the matrix. Such composites can also provide a means of increasing the stiffness of light alloys and of making abrasion-resistant materials. However, the manufacture of MMCs usually involves a stage when the reinforcements are in contact with the molten matrix material. This can lead to chemical interactions and the degradation of the reinforcements, and also to the formation of compounds of intrinsically low resistance to the environment.

The interfacial reaction product Al_4C_3 is a notable example, this compound being water soluble. Another issue with MMCs arises from the fact that the matrix is a conductive metal. If the reinforcement can also conduct current, galvanic couples can form between matrix and reinforcement if these have very different electron energy levels. This has been extensively reviewed by Hihara and Latanision (1994). They propose that corrosion of MMCs in the presence of water and air can be an electrochemical process involving anodic (oxidation) and cathodic (reduction) reactions. This can be written as the dissolution of the metal, M, by the anodic reaction:

$$\mathrm{M} \rightarrow \mathrm{M}^{n+} + n\mathrm{e}^-$$

For the reaction to proceed the electrons released by the anodic reaction must be consumed by a cathodic reaction. In an aqueous environment such reactions can be by oxygen and proton reduction:

$$O_2 + 4e^- + 2H_2O \rightarrow 4OH^-$$

and

$$2H^+ + 2e^- \rightarrow H_2$$

The former reaction requires the presence of oxygen whereas the latter can occur in both aerated environments and in the absence of oxygen. In some cases the oxidation process can lead to the creation of a passive layer on the metal which can eventually inhibit further corrosion, although this protection can be breached if aggressive anions, particularly chlorides, are present. This can result in very severe pitting or dissolution of the metal.

Corrosion by galvanic processes is the major concern in carbon-fiber-reinforced aluminum and magnesium. The fibers, which are electrical conductors, act as relatively inert electrodes for both proton and oxygen reduction. Attempts to limit this behavior by coating the carbon fibers by a variety of materials, such as SiC, B_4C, TiC, and TiB_2, increase the costs of manufacture but work by increasing the electrical resistance between the fiber and the matrix. The protection of the carbon fibers used in such composites may also be necessary as contact between the reinforcements and the molten aluminum leads to the formation of aluminum carbide (Al_4C_3) and decreases in fiber strength. In addition, as mentioned above, the carbide that is formed hydrolyses in damp environments, producing methane (CH_4) and aluminum hydroxide. For some marine applications resin coatings of polyurethane, chlorinated rubber, and epoxy on the outside of the composite have been shown to be effective in preventing this type of corrosion.

Silicon carbide, as particles, fibers, or whiskers, is generally less prone to cause corrosion than carbon fibers; however, it can suffer from galvanic corrosion in aerated chloride-rich environments but this can be controlled using resin coatings.

3. *Ceramic Matrix Composites (CMCs)*

CMCs offer the possibility of structural materials capable of operating at temperatures above 1200 °C in air. The relationship between the fibers and the matrix material is not that found in most composites. The primary role of the fibers is to provide a mechanism to impede crack propagation in the matrix. This is because, although ceramics can withstand high temperatures, they are hard and brittle polycrystalline structures, which usually contain many defects that can initiate crack propagation owing to thermal or mechanical shock. Ceramics do not possess mechanisms for accommodating such shocks so that once a crack is initiated it propagates catastrophically and causes failure of the structure. Introducing fibers into the ceramic provides mechanisms that can absorb the energy of crack propagation. These mechanisms are crack blunting by debonding of the fiber–matrix interface, fiber bridging of cracks in the matrix, and fiber pull-out. These processes rely on there being an interface acting as a mechanical fuse that fails when a crack approaches. In this way CMCs can accumulate multicracking without the structure failing so that they can be used as structural materials, whereas the unreinforced ceramic cannot (Bunsell and Berger 1999).

The most widely studied family of CMCs is composed of β-silicon carbide-based fibers with diameters between 8 μm and 15 μm embedded in a SiC matrix. The fiber–matrix interface is usually controlled by making the surface of the fibers rich in carbon. At room temperature the stress field ahead of a crack in the matrix induces the debonding of the interface with the resulting consequence of an increase in tenacity. Observation of the failure surface reveals extensive fiber pull-out. Use of these CMCs in air at temperatures around 700 °C, however, reveals a dramatic fall in tenacity and no fiber pull-out. Under these conditions, the carbon interphase is oxidized by the oxygen in the air at the interfacial region along the cracks or at surfaces where the interface is revealed. Oxidation proceeds along the interface creating tubular channels and resulting in the oxidation of the fiber surface to create silica (Frety and Boussuge 1990). This silica is produced by the oxidation of nanometer-sized grains of SiC in the fibers. Fibers other than those originally used have near-stoichiometric compositions and as a result have larger SiC grains which are less prone to oxidation, so that overall resistance to degradation is improved. The result of the formation of silica at the interface is to create a strong and brittle bond between the matrix and the fibers, so that instead of debonding when a crack reaches the vicinity of a fiber it passes through it causing it to break with no fiber pull-out (Frety *et al.* 1992). As the mechanisms for absorbing energy have been removed from the composite its tenacity is greatly reduced. The formation of the silica can block the interfacial channels created by the oxidation of the carbon so limiting the degradation. Raising the CMC to temperatures above 1200 °C produces a softening of the silica at the fiber–matrix interface and toughness is increased as debonding and fiber pull-out again become possible.

These composites are produced for applications at very high temperatures in air. First-generation SiC-based fibers begin to lose strength and creep above 1100 °C. This is because of the presence of a disordered phase in the fiber structure which allows creep to occur easily. Temperatures above 1200 °C cause the decomposition of this phase and of the fiber. These problems led to fibers being produced with larger grain sizes which are much more stable. The latest generation of SiC fibers has compositions close to stoichiometry and can retain their properties to around 1400 °C in short-term tests. Eventually these fibers will, however, suffer from oxidation of the SiC and such composites are limited in higher temperature use by this phenomenon.

The development of CMCs made from fibers and matrix materials which are both oxides offers a solution to the limitations of SiC-based fibers. Most oxide fibers are based on alumina, often associated with amorphous mullite. Such fibers suffer from rapid creep at temperatures above 1000 °C and lose their strength by grain boundary weakening above 1100 °C. Fibers produced by 3M under the name Nextel 720 have a two-phase composition of α-alumina and mullite and show good resistance to creep with an upper limit of around 1500 °C. However, the fibers are susceptible to degradation above 1200 °C in the presence of alkaline elements which form silicate phases having low melting points and which consequently promote elemental diffusion in the fibers leading to grain growth and loss of strength.

These issues related to the corrosion resistance of CMCs at elevated temperatures, illustrated above with selected examples, are quite ubiquitous and represent a major issue in their development. The most usual fiber–matrix interfaces that have been developed to toughen these composites by interfacial debonding are based on carbon or boron nitride, both of which have poor resistance to oxidation. This presents a major limitation in their engineering viability (Tressler 1999). Consequently, alternative interfacial phases are being explored such as multilayers which are analogous to the structure of shells of mollusks. Layers in these phases can be of SiC and pyrolytic carbon and, although the carbon is still susceptible to oxidation, the crack deflection behavior that they provide and the increased resistance to oxidation owing to the SiC show possible means of resolving this type of degradation (Naslain *et al.* 1998).

4. Conclusion

Composites of all types can be degraded by the environment in which they are used. This degradation often occurs at the fiber–matrix interface which determines many of the characteristics of this type of material. The environmental resistance of composites varies depending on the material class:

- in polymer matrix composites, water ingress into the composite represents a major issue;
- with MMCs, the presence of matrix/reinforcement reaction products and the possibility of galvanic couple formation between matrix and reinforcement are important issues; and
- with CMCs, the engineering of interfaces which can both toughen the composite by crack deflection and resist attack by oxygen remains a major challenge.

A judicious choice of components and processing conditions, careful engineering of the interface structure, and an understanding of the chemistry underlying the degradation can limit these effects.

Bibliography

Bonniau P, Bunsell A R 1981 A comparative study of water absorption theories applied to glass epoxy composites. *J. Compos. Mater.* **15**, 272–93

Bunsell A R, Berger M -H 1999 *Fine Ceramic Fibers*. Marcel Dekker, New York

Crank J 1956 *The Mathematics of Diffusion*. Clarendon Press, Oxford, UK

Dewimille B, Bunsell A R 1982 The modelling of hydro-thermal ageing in glass-fiber-reinforced epoxy resins. *J. Phys. D* **15**, 2079–91

Frety N, Boussuge M 1990 Relation between high temperature development of fibre–matrix interfaces and the mechanical behaviour of SiC–SiC composites. *Compos. Sci. Technol.* **37**, 177–89

Frety N, Molins R, Boussuge M 1992 Oxidizing ageing effects on SiC–SiC composites. *J. Mater. Sci.* **27**, 5084–90

Hihara L H, Latanision R M 1994 Corrosion of metal matrix composites. *Int. Mater. Rev.* **39** (6), 245–64

Marine R, Rockett T J, Rose V C 1988 Blistering in glass reinforced plastic marine material, Marine Technical Report. University of Rhode Island, Kingston, RI

Naslain R, Pailler R, Bourrat X, Heurtevent F 1998 Mimicking the layered structure of natural shells as a design approach to the fiber–matrix interfaces in CMCs. In: Visconti I C (ed.) *ECCM-8 Science, Technologies and Applications*. Woodhead Publishing, Cambridge, UK, pp. 191–9

Shen C H, Springer G S 1976 Moisture absorption and desorption of composite materials. *J. Compos. Mater.* **10**, 2; **11**, 2–16

Springer G S 1982 Moisture absorption in fiber–resin composites. In: Pritchard G (ed.) *Development in Reinforced Plastics—2*. Applied Science, London, p. 43

Tressler R E 1999 Recent developments in fibres and interphases for high temperature ceramic matrix composites. *Composites A* **30**, 429–37

A. R. Bunsell
Ecole des Mines de Paris, Evry, France

Composites for Biomedical Applications

Implants, and the materials they are made from, have to be biocompatible. Biocompatibility is defined as compatibility of a technical system with the recipient living tissue. One can distinguish between surface and structural compatibility of an implant (Wintermantel *et al.* 1996). Surface compatibility is understood as the chemical, physical, biological, and morphological suitability of the implant surface properties, aiming at clinically desired interactions with the surrounding tissue. Structural compatibility is the optimal adaptation to the mechanical behavior of the hosting tissue, and refers to the mechanical properties of the implant material, such as stiffness and strength. Optimal interaction between biomaterial and host is reached when both surface and structural compatibility requirements are fulfilled.

Many human tissues are anisotropic and reveal a structure similar to that of composite materials. This similarity has been realized, and consequently attempts have been made to use fiber-reinforced composites to replace parts of the human body. Load-bearing implants for bone replacement and fixation made of fiber-reinforced polymers (FRPs) are currently being developed because they offer several advantages over metal implant materials (Wintermantel and Mayer 1995):

- stiffness, strength, and anisotropy can be adapted to the biological requirements over a large range, thus offering optimized structural biocompatibility. Stress shielding can be disabled by adapting the stiffness of the implant material and therefore minimizing strain mismatch between it and bone tissue;
- the absence of metal ions prevents allergic reactions;
- adjusting of x-ray transparency is possible by adding a contrast medium to the polymeric matrix; and
- FRPs cause no artifacts in MRI (magnetic resonance imaging) and CT (computer tomography) relevant to diagnostic interpretation.

Dental composites, on the other hand, make use of reinforcing components, e.g., fibers and particulate filler, mainly to improve their mechanical stability and wear resistance. However, the same basic criteria of biocompatibility, as well as specific advantages of polymer composites towards metallic restoration materials, remain relevant.

1. Fiber-reinforced Implants

Despite the advantages mentioned above, particularly in terms of structural compatibility with bone and other human tissues, composite implants are still mostly at the developmental stage. Competition by alternative materials such as biocompatible metal alloys remains strong, and several critical issues remain to be resolved fully before composite implants can enter in a big way this particularly stringent and competitive application area market segment.

Typical examples of ongoing research in the field of load-bearing implants made of FRPs are hip prostheses, osteosynthesis devices, and intervertebral disk implants. Design, processing, characterization, and clinical application of FRP hip prostheses have been investigated by many researchers (reviewed by Wintermantel and Mayer 1995). The general conclusion of these studies is that stiffness and anisotropy of composite prostheses can be well adapted to cortical bone, thus minimizing strain mismatch and stress shielding effects. The required fatigue strength and surface compatibility considerations, potentially requiring coating of FRP prostheses, are however still a matter of investigation. Some new approaches making use of modified injection molding techniques have also been introduced.

Bone fixation using FRPs has been found to be advantageous because of their low stiffness, thus promoting the formation of external callus, which increases fracture strength after healing (McKibbin 1989). Textile composites, e.g., knitted fabric reinforced composite (KFRC), allow new net-shape processing techniques to be introduced (Mayer and Wintermantel 1998). The textile architecture removes the sensitive interlaminar layers, which induced failure of laminated bone plates. Furthermore, it has been found that formed-in holes increase the fiber density and, therefore, the strength in the critical area around the holes. Further improvements regarding mechanical stability and productivity are achieved using new net-shape flow processes. Their application to bone screws and internal fixation devices is presented below in more detail.

Encouraging studies with intervertebral disk cages consisting of carbon fiber-reinforced polymers (Ciappetta *et al.* 1997) have also been performed. This has led to market introduction of these composite implants.

As with any other material in service, FRPs are subjected to the influence of their environment. The function of the composite part must be fulfilled during its entire lifetime. Certain biomedical applications, e.g., load-bearing implants, are intended to remain in the human body for several decades, where they are exposed to electrolytes.

Unlike metals, FRPs are composed of at least two phases, which introduces different considerations concerning their environmental sensitivity. First, the polymer matrix absorbs moisture, which induces swelling and, in turn, internal stresses. Second, these materials include fiber/matrix interfaces. The mechanical behavior of FRPs, as anisotropic, inhomogeneous materials, is very complex and is a function of the synergistic properties of fiber, matrix, fiber/matrix interface, and of geometric properties such as distribution and orientation of the reinforcement. With regard to environmental degradation, the most critical properties that have to be considered are the interface characteristics. The fiber/matrix interface, as an adhesive bond, is potentially sensitive to moisture degradation. Strength and failure behavior are significantly influenced by the interface, because it is the location of stress transfer.

Many polymeric matrices (Wintermantel and Mayer 1995), such as polysulfone, polyamides, and polyesters, or even epoxy resins, have been applied as matrix materials in preclinical implant development studies. Since elemental carbon is known to have an excellent biocompatibility, carbon–carbon composite

(CCC) systems have also been investigated for implant applications, e.g., heart valves, osteosynthesis devices, and hip endoprostheses stems (Savage 1993). Mainly complex fiber architectures such as three-dimensional textiles are used as reinforcing systems. CCC-implants have to be coated by bioactive coatings such as hydroxyapatite or tricalciumphosphate and because of this they have shown good bone bonding properties. However, the high manufacturing costs were still considered as one of their major drawbacks impeding market introduction at the start of the twenty-first century.

Polyetheretherketone (PEEK) is at the center of interest since this thermoplastic, high-temperature, polymer shows an excellent resistance to most inorganic and organic solvents, to hygrothermal aging, and to radiation. Among the available thermoplastic polymer composite systems, PEEK-carbon fiber composites are among the most intensively investigated systems (Wintermantel and Mayer 1995). Investigations including exposure to physiological solutions and to environments at 37 °C were performed for the evaluation of CF/PEEK in biomedical applications. Summarizing the studies performed on the *in vitro* degradation behavior of CF/PEEK, it is noted that the results are partly in conflict. Some studies report a significant influence on flexural, impact, and interface strengths of CF/PEEK, while other authors note no effects under the same environmental exposure parameters. If effects on mechanical properties were observed, it is generally accepted that only matrix and matrix-dominated properties are affected. Thus, if the occurring stresses in the implant can be carried predominately by the fibers due to appropriate fiber orientation and implant design, one can expect a good long-term stability of the implant in the intracorporeal environment.

Pure PEEK and CF/PEEK have shown no *in vitro* cytotoxicity and have behaved in general comparably to polyethylene (Williams *et al.* 1987). *In vivo* tests indicated nonspecific foreign body reactions around PEEK implants, which are comparable to those of other approved polymeric implants. Carbon fibers could eventually affect the biocompatibility of CF/PEEK implants if fiber particles are released from the implant surface. Therefore, fully matrix covered surfaces have to be achieved, by applying net-shape molding processes for manufacturing of implants and by avoiding any kind of cutting processes. Nevertheless, one has to consider that fretting between implant surface and surrounding tissue will cause the formation of particles if the polymeric surface is not covered by an abrasion-resistant layer, e.g., a spray-coated layer of hydroxyapatite or CVD coatings such as amorphous diamond-like carbon.

Beyond issues related to biocompatibility of composite materials, a specific challenge is in the processing of composite implants. Composite materials show a strong dependence of their microstructure and quality on the specifics of the process by which they are produced (see *Polymer Matrix Composites: Matrices and Processing**). Improvements in the quality and performance of composite implants can, therefore, be achieved by improved and tailored processing. Two examples from current research in the authors' laboratory are given as illustrations of this point.

1.1 Composite Squeeze Forming (CSF)

By using biocompatible, carbon fiber-reinforced PEEK processed with a new composite squeeze forming (CSF) technique, one can optimize the mechanical properties of implants with regard to stress-shielding and fracture fixation stability, with good process flexibility in composite design and cost-effective manufacturing (Wintermantel *et al.* 1998). This process has been used to produce screws, while internal fixation plates have been produced using a variant of the process, called oscillation CSF. This process provides more homogeneous mechanical properties and reinforces sensitive areas, e.g., the countersunk hole. The imaging contrast behavior for x-ray based systems was optimized using various types of metal filaments incorporated in the composite in the range of 0.5–1.0 vol.%. The design concept of the internal fixation system aims at the lowest possible irritation of the fractured site. This was achieved through a hybrid design featuring characteristics of bone plate as well as internal fixators. Bone screws made by transfer CSF achieved tensile strength of 400 MPa and bending strength of 700–1000 MPa. Fixation plates manufactured by oscillation CSF had maximum bending moments of 14 Nm, thus equaling their counterparts made of commercially pure (CP)-Titanium (13 Nm). Figure 1 illustrates the synergistic effect of process and component shape optimization on the fatigue properties of the screw being applied in the mechanically most critical case of nonunion of transvertebral fixation. It is seen that through process and design improvements, composite implant performance can be rendered competitive with that of metallic screws.

1.2 Injection Molding

As an example of a highly loaded implant, an anisotropic hip endoprosthesis stem can be used to demonstrate the potential of the combination of an optimized material based on injection molded short carbon fiber-reinforced PEEK with a new processing technique (Semadeni *et al.* 1998). For this application, the fiber content must be high. However, with increasing fiber content one observes a pronounced deterioration of the fibers after processing, down to a residual length of 50 μm at 68 wt.%. Analysis of the influence of fiber content and length on the structure–properties

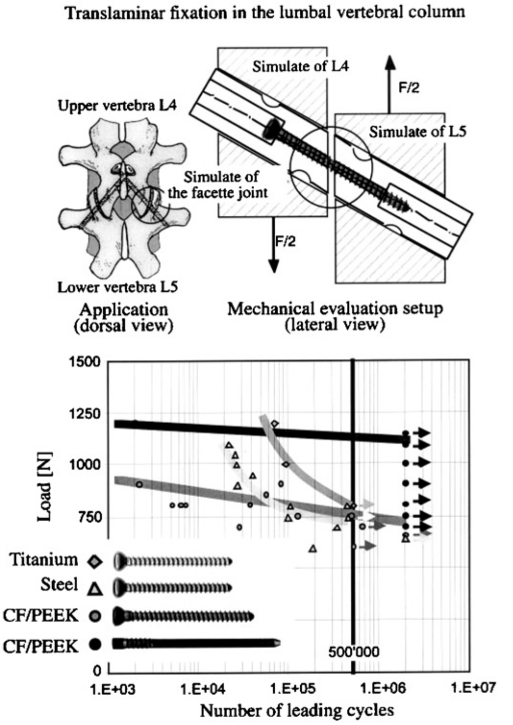


Figure 1
Fatigue properties of titanium and carbon fiber-reinforced bone screws illustrating the synergistic effect of process and design optimization. The load case is mainly transversal shear and short bending is indicated in the sketch.

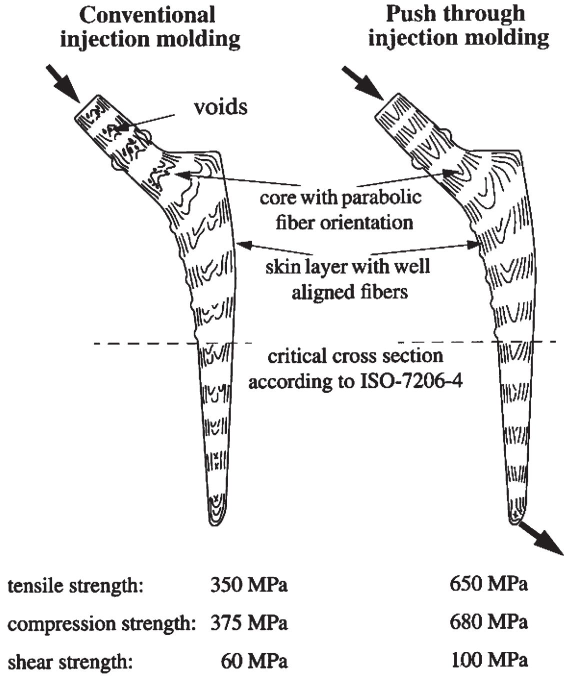


Figure 2
Influence of the processing technique and fiber content on the specific mechanical properties of the prosthesis in its critical cross-section: comparison of a conventionally injection molded prosthesis (left, 30 wt.% fiber content) with a push-through molded prosthesis (right, 50 wt.% fiber content). Strength values take reference to the ISO-7206-4 standard for distal fixation of stems.

relations in the investigated tensile test specimens revealed that the highest mechanical strength (700 MPa) is achieved at 50 wt.% fiber. Above this value, structural inhomogeneities, e.g., domains and localized vortices, are strength limiting. Conventional injection molding and push–pull molding were compared to new processing techniques, i.e., push-out and push-through molding. The working principle of these new injection molding processes is the maintenance of a continuous flow in the part during consolidation, thereby providing process-integrated control over the fiber orientation distribution. The evaluation of this specific potential in comparison to push–pull molding and conventional injection molding indicates that the new technique provides additional mechanical performance if fiber content and fiber length are increased and when the process is applied to bulky parts, such as the hip endoprosthesis stem. Compared to commercially available carbon fiber-reinforced PEEK (Victrex 450CA30) processed by conventional injection molding, the combination of the optimized material (APC2/IM7, 50 wt.% fiber content) with push-through injection molding allows to increase the modulus and strength by 171% and 80%, respectively. This can be closely related to the homogeneity of the consolidation, the fiber orientation distribution, which exhibits a process-induced extension of the highly oriented surface layer, and, finally, the increase in residual fiber length (500 μm instead of 100 μm), see Fig. 2. Clinically even more relevant, a ratio of fatigue to static strength of 0.69 is obtained.

These two examples indicate the importance of processing towards the development of competitive composite implant technology.

2. *Dental Composites*

Auto- and light-cured composites are used for direct and indirect restorations in dentistry. Initially, composites were bonded to enamel only (Buonocore 1955). Since bonding to dentin was established, composites became a standard restorative material in western industrial nations (Lutz *et al.* 1993). Dental composites generally comprise three phases: (i) an organic phase, the matrix; (ii) a dispersed phase, the fillers, e.g., particles or fibers; and (iii) coupling agents, the interfacial phase (Vanherle and Smith 1985). A concise review covering more than 300 publications in the field of dental restoration materials, which includes dental composites, is given by Whitters *et al.* (1999).

The common monomers are based on diacrylates of the formula M–R–M or oligoacrylate of formula M–R–M (n) are used. R is the skeleton of the molecule and M is the methacrylic group –O–CO–$C(CH_3) = CH_2$, which includes a double bond for polymerization. The most important representatives for R are polycarbonate, polyurethane, and polysiloxane, the latter and most recent also known as ormoceres.

Polymers are built up more or less completely in homo- or co-polymerization, network formation and interlinking. High conversion rates of double bonds after polymerization are desired for complete polymerization. Unreacted double bonds can result in material softening and higher water absorption, which can lead to an unwanted change of color and, even more important, can cause leaching of toxic components and degradation of mechanical properties. Activated double bonds with free radicals on free surfaces are useful for co-polymerization between composite-layers. The residual monomer content should be kept to a minimum, because of its allergic and toxicologic potential.

2.1 Filler Systems

The available particulate filler systems can be subdivided according to filler size and properties:

(*a*) *Microfillers and microfiller-complexes.* SiO_2-microfillers of sizes ranging from 5 nm to 1000 nm with an average of about 200 nm have the problem of large active free surfaces, which makes their incorporation into the matrix difficult, limiting possible filler contents in the material. The technical solution is the use of larger and/or spherical microfillers of 200–300 nm in diameter, or the use of microfiller complexes. Splintered pre-polymerized microfiller complexes are made by incorporating microfillers into an organic resin matrix and milling the cured mixture into particles of diameter near 1 μm. The surface of these complexes is not very reactive to the matrix. Naturally or artificially agglomerated microfiller complexes are also 1 μm in diameter and almost purely inorganic. Their surface might be mechanically retentive. Spherical pre-polymerized complexes are made by mixing matrix and microfiller, forming spheres of about 1 μm in diameter. Complex size distributions allow extremely dense packing at convenient viscosity. Bonding to the matrix is rather weak.

(*b*) *Solid and porous macrofillers.* Solid macrofillers are prepared by grinding or crushing of inorganic materials like quartz, glass, borosilicate, or ceramics to sizes ideally in the range 0.8–2 μm. Heavy-metal glasses provide clinically desired radioopacity. Porous macrofillers are prepared by etching or sintering solid fillers of diameter up to 20 μm and 30 μm. Traditional macrofillers can be silanated or impregnated with monomer.

(*c*) *Fibrous fillers.* Besides glass and silica fibers, various fiber systems, e.g., carbon fibers, silicon nitride, all being in the range of 5–10 μm in diameter and 60–100 μm in length, are being investigated. Silicon nitride whiskers (diameter 0.4 μm, length 5 μm) have been introduced, which are roughened with fused silica to enhance the interfacial bonding to the polymeric matrix. Due to their high aspect ratio fibrous filler induce anisotropy in the composite and rather high residual stresses in the cured resin.

Also in dental composites, the interfacial properties are of key importance for the long-term performance of the material. A good interfacial phase can be achieved by silanating or silicating of the inorganic fillers or by co-polymerization between the matrix and combined organic and inorganic fillers. All current versions of interface adhesion are still sensitive to chemical disintegration. Additionally, mechanical interlocking is attempted with natural and artificial agglomerated microfiller complexes and splintered prepolymerized microfiller complexes. The highest degree of interlocking is achieved in porous macrofillers.

2.2 Classification and Clinical Performance of Dental Composites

The types of fillers previously described represent the backbone of the classification of composites (Lutz and Phillips 1983).

(*a*) *Homogeneous microfilled composites.* Homogeneous microfilled composites are constructed from a matrix and microfillers of a diameter of 400 nm. They feature low filler content, have a low elastic modulus, and are not radioopaque. Their surface is smooth, easy to polish, and shiny. Wear of the restoration and of the antagonistic teeth is low. However, it is deformable and has a low form stability.

(*b*) *Inhomogeneous microfilled composites.* Inhomogeneous microfilled composites are constructed from a matrix, microfillers, and microfiller complexes. They feature moderate filler content, have a low elastic modulus, and their radioopacity is acceptable. Their surface is also smooth, easy to polish, and shiny. Wear of the restoration and the antagonists is low, but it is deformable and has a low form stability. This confines their use to non stress-bearing restorations.

(*c*) *Traditional composites*. Traditional composites are constructed from a matrix and solid and/or porous fillers. They feature high filler content and high hardness, their Young's modulus is dentin-like, and their radioopacity is good. Depending on the degree of milling, the surface is rough to nearly smooth. Depending on size and distribution of the fillers, wear of the restoration and antagonists is low to high.

(*d*) *Hybrid composites*. Hybrid composites are constructed from a matrix, microfillers, traditional fillers, and rarely from glass fibers. Depending on the size of the traditional fillers, these can be divided into coarse and fine hybrid composites. They are always highly filled, their Young's modulus is dentin-like, and the radioopacity is good. In coarse hybrid composites the traditional fillers are about 3–5 μm and their characteristics are comparable to those of traditional composites. In fine hybrid composites the size of the traditional fillers is 2 μm. The surface is nearly smooth, acceptable to polish, and wear of the restoration and the antagonists can be equal to those of human enamel. The stability of the restoration is good.

2.3 *Prospects*

Dental composites are hydrophobic in relation to human dental tissues. To achieve bonding to enamel and dentin, a surface pretreatment is necessary. They are still sensitive to humidity during processing in the oral cavity. The shrinkage of modern dental composite materials amounts to 2–3 vol.%. During polymerization the shrinkage forces can reach 3–7 MPa with peaks of 30 MPa. For these reasons, an advanced operative technique is necessary. A shrinkage-free composite with a good wetting ability is highly desired to simplify the operative technique and is the object of ongoing research.

Glass and Kevlar continuous fiber-reinforced thermoplastics are being developed for specific applications, e.g., prosthodontic frameworks and orthodontic retainers, however, these still require considerable improvement of the interfacial stability. To explore possible enhancement of the mechanical properties of resin composites, strong electric fields are investigated inducing an alignment of filler particles and, thus, increasing the mechanical properties of filler-reinforced systems.

The hydrolytic sensitivity of the available systems compromising mechanical strength and fracture toughness remains an unsolved problem that has to be approached.

3. *Conclusion*

Composite materials are attractive in biomaterial applications because these can mimic natural load transmitting structures, e.g., bone, anisotropic reinforcements for osteosynthesis and for hip endoprosthesis stems. It is possible to process carbon fibers and thermoplastic matrices into composites which fulfill the following requirements:

(i) optimized interface and interphase designs for long term mechanical stability;

(ii) mimicking bone anisotropy with three dimensional reinforcement architecture due to appropriate processing; and

(iii) cost-effective industrial manufacturing induced by net shape processing.

Dental composites are now well established, yet undergo a continuous development aiming at reduction of shrinkage and wear as well as improvement of mechanical long term stability. Since amalgams have still to be considered as a strong benchmark and since composites are still relatively new to implant applications further developments are required. These are in progress integrating the advances in nanocomposites and thermoplastic high performance polymers.

Bibliography

Buonocore M G 1955 A simple method of increasing the adhesion of acrylic filling materials to enamel surfaces. *J. Dent. Res.* **34**, 849–53

Ciappetta P, Boriani S, Fava G 1997 A carbon fiber-reinforced polymer cage for vertebral body replacement: technical note. *Neurosurgery* **41**, 1203–6

Lutz F, Krejci I, Schüpbach P 1993 Adhäsivsysteme für zahnfarbene restaurationen. *Schweiz. Monatsschr. Zahnmed.* **103**, 537–49

Lutz F, Phillips R W 1983 A classification and evaluation of composite resin systems. *J. Prosthet. Dent.* **50**, 480–8

Mayer J, Wintermantel E 1998 Thermoforming processes for knitted fiber reinforced thermoplastics: new manufacturing techniques for load bearing, anisotropic implants. In: Bhattacharyya D (ed.) *Composite Sheet Forming*. Elsevier, Amsterdam, pp. 403–40

McKibbin B 1989 Biological considerations in osteosynthesis. *Chirurgie* **115**, 683–6

Savage G 1993 In: *Carbon–Carbon Composites*. Chapman and Hall, London, p. 156

Semadeni M, Zerlik H, Mayer J, Wintermantel E 1998 Processing-properties relations in injection molded PEEK reinforced with up to 60 vol.% carbon fibers. *Polym. Polym. Compos.* **6**, 279–86

Vanherle G, Smith D C 1985 In: *Posterior Composite Resin Dental Restorative Materials*. Peter Szulc Publishing Co, Utrecht, The Netherlands, pp. 109–60

Whitters C J, Strang R, Brown D, Clarke R L, Curtis R V, Hatton P V, Ireland A J, Lloyd C H, McGabe J F, Nicholson J W, Scrimgeour S N, Sectos J C, Sherriff M, von Noort R, Watts D C, Wood D 1999 Dental materials: 1997 literature review. *J. Dentistry* **27**, 401–35

Williams D F, McNamara A, Turner R M 1987 Potential of polyetheretherketone (PEEK) and carbon-fiber-reinforced PEEL in medical applications. *J. Mater. Sci. Lett.* **6**, 188–90

Wintermantel E, Bruinink A, Eckert K L, Ruffieux K, Petitmermet M, Mayer J 1998 Tissue engineering supported with structured biocompatible materials; goals and achievements.

In: Speidel M O, Uggowitzer P J (eds.) *Materials in Medicine*. VDF Verlag, Zurich, Switzerland, pp. 1–136

Wintermantel E, Mayer J 1995 Anisotropic biomaterials: strategies and developments for bone implants. In: Wise D L (ed.) *Encyclopedic Handbook of Biomaterials and Bioengineering*. Dekker, New York, Vol. B, pp. 3–42

Wintermantel E, Mayer J, Blum J, Eckert K L, Ruffieux K 1996 Tissue engineering scaffolds using superstructures. *Biomaterials* **17**, 83–8

E. Wintermantel[a], J. Mayer[b] and T. N. Goehring[c]
[a] *Technical University Munich, Germany*
[b] *Swiss Federal Institute of Technology, Schlieren, Switzerland*
[c] *University of Zurich, Germany*

Composites for Sensors and Actuators

Composite materials by nature have the advantage of allowing the combination of significantly different types of material. This has principally been used to integrate reinforcing fibers with a passive load-bearing functionality, which carry load along selected and optimized paths. This allows one mainly to avoid fatigue problems and to optimize strength and weight of structural components. Composites can, however, also integrate reinforcing elements that feature other functionalities than load-bearing ones, such as sensing and/or actuation. Resulting materials and structures are known as *adaptive, intelligent*, or *smart*.

Smart materials and structures—briefly expressed—are materials and structures with built-in sensing and/or actuation functions, which contain sensors and actuators linked via a controller. The level of intimacy of coupling between the material and its active elements can range from a classical mechanical coupling of two structures (such as is found in many mechanical structures) down to the level of molecular design. As such, "smart" composites represent a convenient intermediate level of coupling, which allows one to conveniently design and develop a smart material to be applied in a smart structure.

The reasons for developing smart materials and structures are various. These range from a desire to sense the material condition in service (e.g., damage, loads, shape, temperature, pressure, etc.) to design materials capable of adaptive actuation (e.g., changing deformation, shape, stiffness, inner stresses, temperature, color, etc.). Controllers may be built into the composite itself, but so far these have been situated outside the material in the vast majority of cases. The matrix material of such composites is mainly polymer resin; however, metal or concrete matrices have also been considered.

Electromagnetic composites represent another type of smart composite allowing one to modify dielectric permittivity, magnetic permeability, and electric resistivity, which is used for stealth applications. These composites are a combination of multilayered structures, or feature dispersed particles in a matrix. The particles used for passive absorbers are typically graphite powder/fibers, aluminum particles, iron whiskers, or nickel flakes. Active absorbers only absorb electromagnetic energy through integrated semiconductor diodes or other electronic devices when being triggered externally. Further embedded materials can be photosensitive (e.g., fiber optics), pyrosensitive (e.g., silver iodide, Ni–Cr filaments), or voltage/current sensitive arrays of elements, which can be designed for either narrow- or broad-band frequency selection.

With respect to different frequency bands (radio frequency to microwaves), further active roles for "smart" materials can be designed. Examples include phantom materials (absorbing composites simulating human bone tissue and body fluid), electromagnetic shielding (lossy dielectric layers), quasioptic composites consisting of rare-earth materials and ferrites, composite dielectrics (which are two-phase mixtures of low- and high-permittivity components such as polystyrene and barium titanate), chiral composites which include particulates of asymmetric shape allowing the cross-coupling of electric and magnetic effects, and electrochemical composites such as porous carbon impregnated with hydrophobic agents used for fuel cells.

Techniques related to nanoscale materials and composites are atomic and molecular layer manipulation with a focus on semiconductors.

1. Sensor, Actuator, and Control Options

1.1 Sensor Options

Components already existing in a classical structural composite may be used as sensors. An example is carbon fibers, of which variations in the electrical resistance can be used for monitoring damage such as delaminations, or matrix or fiber cracking. A wider application can be seen with fiber optical sensors being attached to, or integrated into, conventional structural composites. Polyimide-clad Bragg grating sensors are the most popular type of such composites. Another option is to make use of piezoelectric materials, available as both ceramics and polymers. These emit a voltage when subjected to stress, and can be shaped to any size and form and adapted onto, or integrated into, the composite. Furthermore, piezoelectric elements can additionally work as actuators when a voltage is applied. Finally, the relationship between electrical resistance and strain of shape memory alloys (SMAs) is another option allowing one to use a classical actuation element as a sensor as well.

1.2 Actuator Options

Piezoelectric materials, usually ceramics (e.g., lead–zirconate–titanate, PZT), are highly favored as actuators owing to their high precision and very rapid response; however, piezoelectric actuators can only work under dynamic loading conditions. Patches or even fibers can be integrated into, or bonded onto, the composite. Deflections achieved are relatively small and may be used for influencing structural damping behavior. Larger deflections need to be achieved through leverage mechanisms. Brittleness of the piezoelectric ceramics, manifested as low allowable strain, must be considered in these applications. Magnetostrictive materials have similar or even slightly better properties as actuators when compared to piezoelectric materials; here the electrical field is replaced by a magnetic field.

SMAs allow one to recover large deformations as a result of a transformation process between martensite and austenite and work at frequencies of about 3 Hz at maximum. Their change in Young's modulus by a factor of three between martensite and austenite allows one to influence damping properties. Transformation is, however, mainly achieved at temperatures less than 60 °C, which limits their applications.

The shape memory effect is also observed for specific classes of polymers (e.g., polyurethane of the polyester polyole series) but polymeric materials produce about two orders of magnitude less in force when compared to more conventional metallic SMAs. Density is much lower ($0.9–1.1\,\mathrm{g\,cm^{-3}}$ vs. $6–8\,\mathrm{g\,cm^{-3}}$) and deformation significantly higher (250–800% vs. 6–8%). More information on these materials can be obtained from the literature (e.g., Otsuka and Wayman 1998).

Other types of polymeric composite worth mentioning are ionic polymer–metal composites (IPMCs). These are mainly used as artificial muscles based on charged polyelectrolyte membranes such as collagen or fibrous protein, poly(acrylic acid) (PAA), and poly(vinyl chloride) (PVC). IPMCs show large deformation ($>10\%$) in the presence of low applied voltage (4–7 V) and operate best in a humid environment, but can also be made as self-contained encapsulated actuators to operate in dry environments. Their density is similar to that of the polymers mentioned above but the reaction time is of the order of a microsecond to a second, and thus these are faster. More details are given in Shahinpoor *et al.* (1998).

Electrorheological (ER) fluids are used in order to influence the damping characteristics of materials. These consist of a dielectric liquid (e.g., oil), polarizable particles (e.g., cornstarch), and water as an activator. Through application of an electrical field their rheological properties (e.g., viscosity) are changed. In a zero-field environment the particles are randomly distributed. In an electrical field, however, the particles line up according to the field lines, which then leads to the more rigid behavior of the fluid.

1.3 Control Options

The classical procedures of feedforward and feedback control are used for smart structures as well. Owing to the possibly large number of sensors and actuators to be handled, artificial neural networks have gained broad interest as a means for pattern recognition and novelty detection. Genetic algorithms and wavelet transforms are further procedures that allow for automated determination of optimum solutions and advanced sensor signal processing, respectively.

2. Design and Manufacturing

2.1 General Considerations

A major consideration is the compatibility between the sensing/actuation and the host materials. The stiffness ratio between these different materials is a significant parameter. If the actuation or sensing material is too soft, then it will be difficult to transfer the actuation to the host material for actuation, or to the sensing material for sensing designs. A value for this stiffness ratio between 0.5 and 2 is a good design rule of thumb. Another important issue is the quality of the interface between the different materials, because this must have the ability to transfer significant shear stresses. Otherwise the sensing/actuation elements work rather as artificial delaminations. Poor interface stress transfer can also result from chemical incompatibility, or from increased temperature generated by heating of the actuation element.

Size, shape, and quantity of the sensing/actuation elements need to be considered with regard to any notch effects these elements may cause when being integrated into the composite (Warkentin and Crawley 1991). Depending on the application, a large number of widely dispersed small elements may be more useful than one bulky element placed at one location only. The location and thus the number of sensing/actuation elements also very much depend on the geometry and function of the component and structure considered. Optimum design of the smart composite can therefore often very much rely on global component design.

Smart composites are not manufactured on an industrial scale basis, nor are such materials yet qualified according to industrial standards. The vast majority of activity on "smart" composites and structures is still at a research and development stage. There are, however, practical applications for which research and development of such composites has achieved relatively advanced levels, and in which their use is generating significant interest. A few of these cases are described below.

2.2 Fiber Optics

Optical fibers are among the most frequently used sensing elements in "smart composites." These are preferably embedded in the composite (Fig. 1), because this generates a more reliable sensor signal output; however, there are techniques that allow surface attachment while maintaining sensor output quality. The smaller the diameter of the optical fiber, the better is its integration into the composite, with diameters ⩽125 μm being ideal. Larger diameters are feasible; however, one needs to take into account the formation of resin-rich zones around the fiber. These zones are especially likely if the optical fibers are embedded at an angle to the adjacent ply directions. Very thin optical fibers may, however, lead to positioning problems during processing owing to electrostatic forces. Cleaning of the optical fibers with acetone helps to improve adhesion significantly. One advantage of the use of optical fibers as sensors is that structural integrity is not compromised, since the volume fraction of optical fibers usually required is small compared to the volume fraction of the reinforcing fibers. Another advantage of fiber optic sensors is their very low weight and volume and their immunity to electromagnetic interference.

Critical locations of fracture in the resulting composites are the intersections where the optical fiber comes out of the material, as well as optical fiber terminations and connections. A solution to this makes use of locally cocured synthetic rubber implants that are put around the fiber at these intersections. Adjacent to these are connectors, which are conventional optical fiber plugs incorporating a ceramic precision ferrule. Special techniques have been developed to integrate these connectors into the composite. Results of some initial studies are reported in Green *et al.* (2000).

2.3 Piezoelectric Materials

Piezoelectric sensors are kept small (a few millimeters in diameter and less than 1 mm thick). Examples are "smart layers" (http://www.acellent.net) made of dielectric films with an embedded network of distributed piezoceramic elements, manufactured by a process similar to a circuit printing technique. The layer can either be integrated into the prepreg composite, or it can be bonded onto its surface without compromising the strength of the composite.

The use of piezoelectric elements as actuators (Fig. 2(a)) requires elements of relatively large size, such as 10 mm × 4 mm × 0.1 mm. Special cutouts have to be prepared in the prepreg material in order to integrate such actuation devices. Considerable care

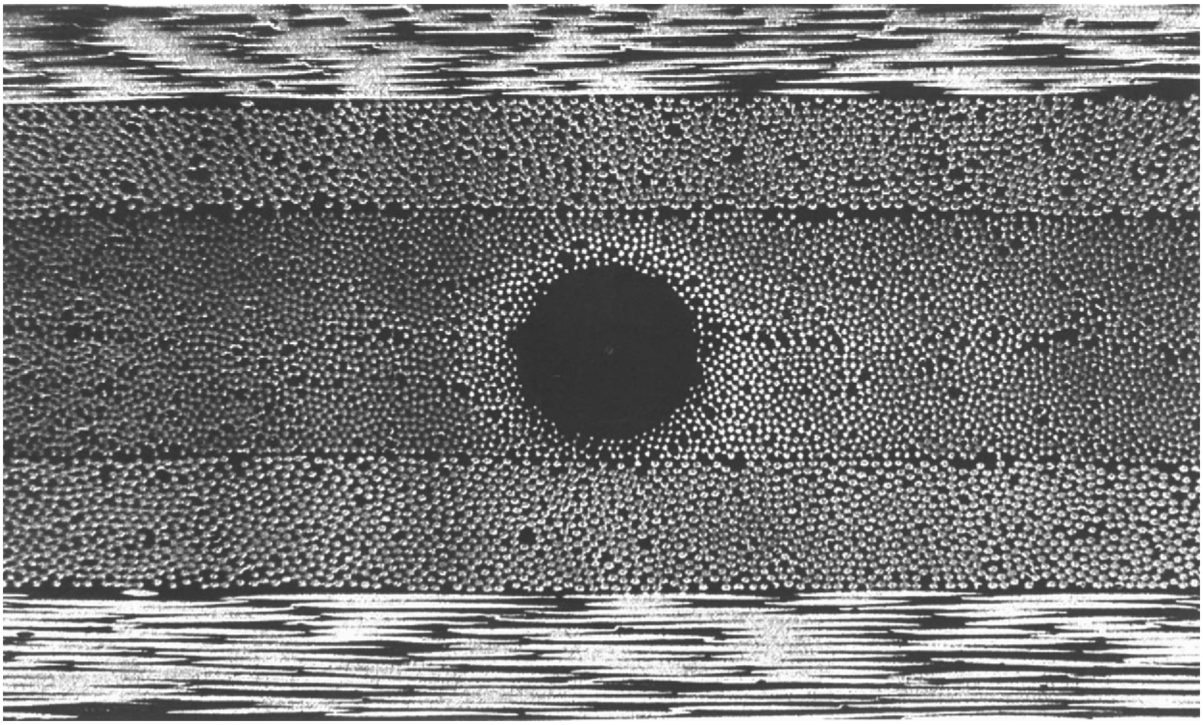
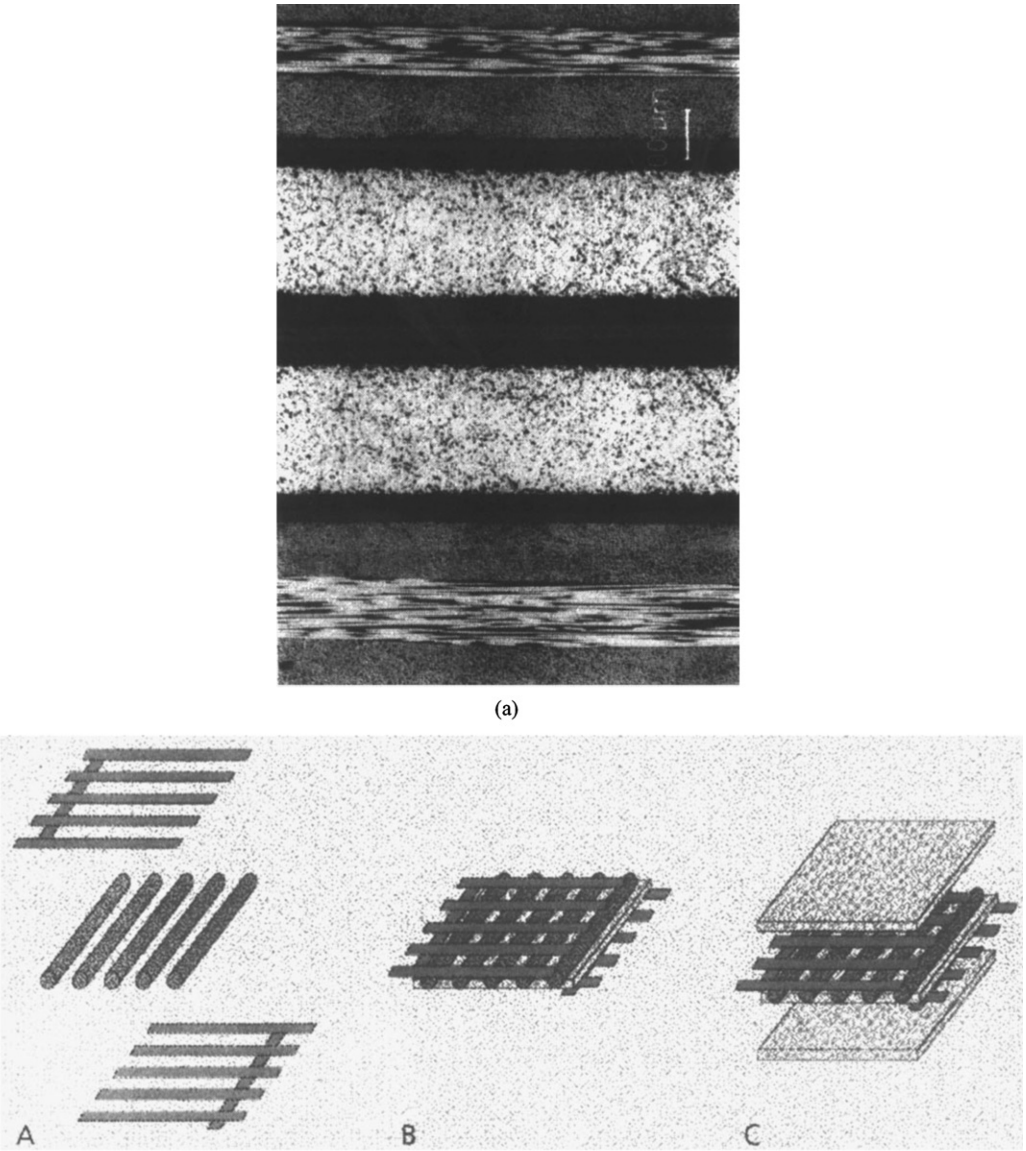


Figure 1
Fiber optic sensor in a quasi-isotropic laminate (courtesy of M. Trutzel, DaimlerChrysler R&T).



(a)

(b)

Figure 2
(a) Piezoelectric wafers laminated into a composite (courtesy of H. Zaglauer, U. Herold-Schmidt, F. Dürr, DaimlerChrysler R&T). (b) Piezoelectric fiber concept for integration into composites (courtesy of H. Schönecker *et al.*, Fraunhofer-Institute, IKTS, Dresden, Germany).

is required regarding the selection of the appropriate bonding technique to avoid any detrimental effect to the composite. A compressive prestress on the piezoelectric element can help to improve its performance. When integrating the piezoelectric elements care is also required in order to limit the influence of heat being generated by these actuators when operating at higher frequencies and possibly as an actuator cluster. Special conductive layers for heat transfer help to alleviate this problem.

Piezoelectric fibers (Fig. 2(b)) can be integrated like conventional fibers and allow one to avoid any cutouts in the prepregs. A problem, however, exists in producing the required electrical field; this has been solved by developing an interdigitated electrode pattern that is put on the top and bottom of a layer of piezoelectric fibers. This new type of prepreg can then be included into the stacking sequence during composite production.

Other types of piezoelectric elements used as actuators are 1–3 piezoelectric actuator composites. This type of composite consists of a series of piezoelectric rods surrounded by a soft matrix. The rods are manufactured by an injection molding process and are directed vertically to the plane of the composite, which is also the direction of polarization. This allows one to make the best use of sensing and actuation capabilities of the piezoelectric elements, as long as deformations are required in the thickness direction of a plate.

To overcome the problem of high brittleness of piezoelectric ceramics, piezoelectric rubbers and paints have been developed based on the principle that a piezoelectric material is milled into a very fine particulate-sized powder, mixed to a rubber or epoxy resin, and then polarized to obtain again the piezoelectric effect. This allows one easily to shape these types of sensors and actuators. However, a reduction in the sensing and actuation effect is observed when compared to the original piezoelectric material.

2.4 SMAs

State-of-the-art integration of SMA elements—mainly wires with films as an option—into a composite material requires a frame with a comb at its ends where the SMA wire is continuously looped around (Fig. 3). The ends of the frame in the SMA fiber direction are then pulled, which allows one to produce a desired prestress and prestrain on the wires. The whole system is then stacked with the other prepregs and cured, producing a SMA reinforced composite. Appropriate surface treatment of the wires plays a major role in allowing shear stresses to be transferred; however, this should not compromise strength. Techniques of weaving SMA wires into prepregs and even designing complete preforms are a further step being explored. An important issue is that of heat generated by the SMA when going through larger deformations and higher frequencies, and which needs to be transferred from the interface between the SMA and host materials.

SMAs may also be plasma sprayed or sputtered as very thin coatings of arbitrary shape on either the composite surface or the prepreg surface prior to composite consolidation. These developments are, however, still at a very early stage. Another novel approach is the integration of SMA particles into a metal matrix, which leads to an arrest effect on crack propagation.

2.5 ER Fluids

ER fluids by definition can only be filled in cavities between composite plates and rubber seals along the thickness axis. CFRP plates are advantageous because they can be used directly as electrodes owing to their electrical conductivity.

2.6 Electrochromatic Effects

Electrochromatic windows allow one to change light transmission as a result of applied voltage. The window consists of a sandwiched cross-section including two transparent electrodes, two electrochromatic layers, and an electrolytic polymer layer. A small current of 1–2 V imposed by the electrodes on the system allows an exchange of ions between the two electrochromatic layers owing to a chemical process and leads to darkening of these layers. Reversing the direction of the current allows one to reverse the darkening process.

2.7 Manufacturing Methods

Manufacturing of composites for sensors and actuators requires care because the sensors, actuators, and wiring must not be affected by mechanical, thermal, or chemical reactions during processing. This is with respect to their properties as well as location within the composite. Adequate care may be achieved through individual hand laying and bonding; however, this is the most costly process. Initial research activities are under way to integrate active fibers such as the ones made of SMA or piezoelectric materials using automatic methods. Automated fiber and tape laying are alternatives but these techniques still need to be adapted to the specifics of these new composites. One approach is a filament winding/tape laying technique with directed energy heating through a laser to synthesize a composite structure layer by layer. Finally, coating techniques represent another promising area where homogeneous material qualities have, however, to be guaranteed. Other low-cost manufacturing processes that have been proposed include extrusion processes for polymer materials designed such that optical fibers or films of piezoelectric materials can be included in the material.

3. Applications

3.1 Sensing

Monitoring of loads is a big issue for determining used structural life in components. The sensing elements that are mainly considered are fiber optics and piezoelectric elements, the latter being useful for vibrational loads of frequencies >10Hz. A related issue is damage monitoring, which includes detection of barely visible delaminations, or of localized fiber and matrix fracture resulting from overloads such as impacts. Here fiber optic sensors are mainly used for monitoring damage-inducing loads while the

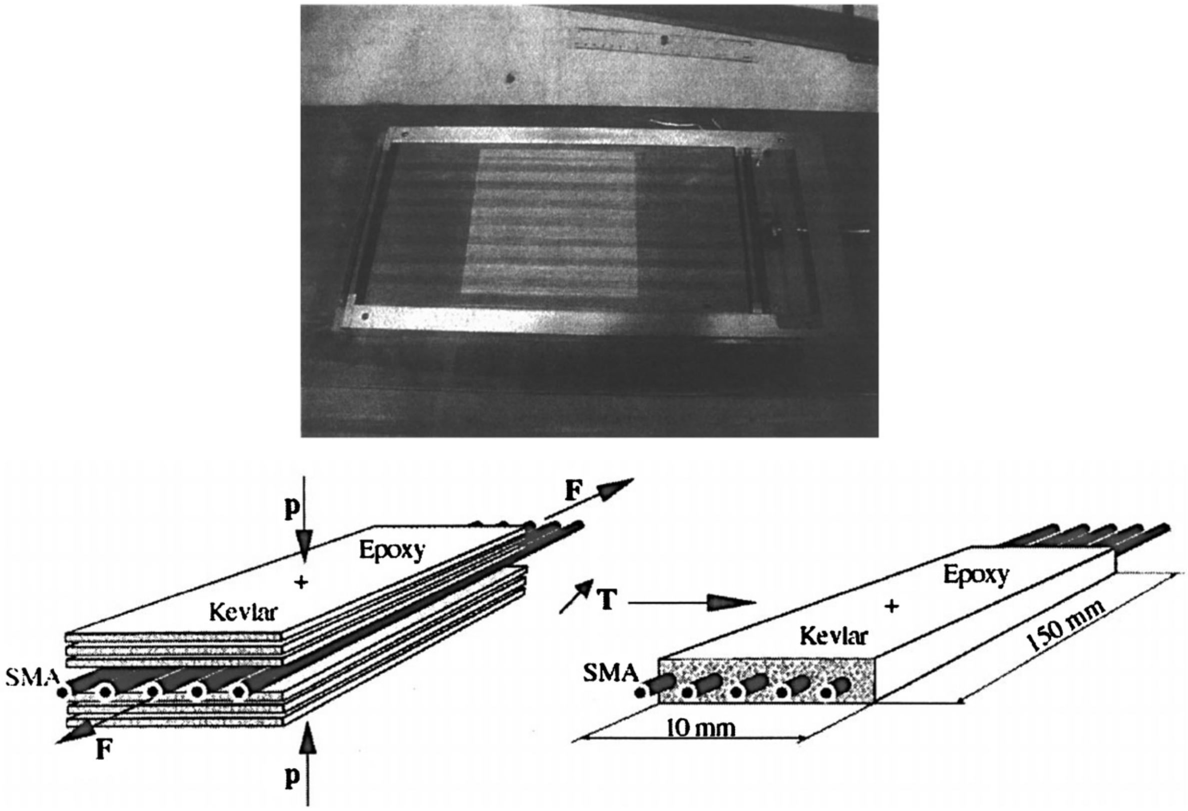


Figure 3
Manufacturing of SMA-reinforced composites. Frame for prestraining SMA fibers (top), stacking and curing process (bottom) (courtesy of V. Michaud, R. Gotthardt *et al.*, EPFL Lausanne, Switzerland).

piezoelectric elements can even be used as actuators by sending into the material a chosen (e.g., ultrasonic) signal which is then recorded by another sensor. Nondestructive testing thus becomes an integral part of the composite material. Another option specific for carbon fiber-reinforced composites is to monitor electrical resistance. This can be achieved by simply contacting the carbon fibers with an adequate current generating and sensing device across the composite. Fiber optics and piezoelectric sensors are also well suited for monitoring the curing process during composite processing. Finally, one can mention electromagnetic sensing, which can be conducted using specially designed composites, and is classically known for radars in aerospace and general defense applications, but is becoming increasingly popular with ground traffic control as well.

3.2 Actuation

Apart from introducing ultrasonic signals, shape control, vibration damping, and electromagnetic antennas are further significant applications of "smart" composites. Shape control originally comes from a desire to adapt the shape of fluid dynamic profiles according to different operational conditions by changing their camber or twisting rotorblades. Shape control can, however, also be desired in order to generate or control the grips of arms such as those used for robots. Applications can also be seen in the medical field, such as stents or valves. The actuators mainly considered are ones that make use of SMAs and, for some specific applications such as microrobots with very low forces, also shape memory polymers. For vibration damping applications, improved passive damping can be achieved by integrating SMAs into the composite, with the side effect that the dynamic properties of the SMA can be changed by heating. SMA-reinforced composites are also used because of their improved impact resistance when compared to the standard composite. Active damping can be added to other functionalities by using piezoelectric elements and/or ER fluids. Much effort is focused on active constrained layer damping using sandwiched structures consisting of the parent structure, a viscoelastic layer, a constraining layer, and a

PZT actuator. The actuator is then able to distribute more efficiently its damping energy through the constrained layer over a wider area. Specific applications are seen with turboengines as well as in aircraft generating noise. 1–3 piezoelectric composites represent another solution used as hydrophones for acoustic damping. Actuation is also seen with electromagnetic wave-absorbing composites for antennas, which are used to send out well-defined signals but can be converted to materials that are electromechanically invisible with regard to stealth applications.

4. Conclusion

Although composites with sensors and actuators are generally not available "off the shelf," their design as well as their potential applications are already well understood to a large extent. Full-scale experimental verification in an application-oriented environment will allow one to identify and hopefully close any remaining gaps. Specification, qualification, and manufacturing procedures can then be established that may have an impact on further improvement of available sensor and actuator materials.

Bibliography

Green A K, Zaidman M, Shafir E, Tur M, Gali S 2000 Infrastructure development for incorporating fibre-optic sensors in composite materials. *Smart Mater. Struct.* **9**, 316–21

Otsuka K, Wayman C M (eds.) 1998 *Shape Memory Materials*. Cambridge University Press, Cambridge

Shahinpoor M, Bar-Cohen Y, Simpson J O, Smith J 1998 Ionic polymer–metal composites (IPMCs) as biomimetic sensors, actuators and artificial muscles—a review. *Smart Mater. Struct.* **7**, R15–30

Warkentin D J, Crawley E F 1991 Embedded electronics for intelligent structures. *Proc. AIAA Structures, Structural Dynamics and Materials Conf.*, AIAA-91-1084-CP. AIAA, Baltimore, MD

C. Boller
European Aeronautic Defence and Space Company,
Muenchen, Germany

Composites, Joining of

Composite materials, be they metal matrix composites (MMCs), polymer matrix composites (PMCs), or ceramic matrix composites (CMCs), are joined by the same procedures that are well-known for joining isotropic materials such as metals or polymers. This is to be expected since the process of transferring loads between joined components sets up stress patterns which are more dependent on the component/joint configuration than on the material used to fabricate the article.

Of course, composites often display directionality of mechanical properties, especially those fabricated from long or continuous fibers and used in laminated form e.g., graphite fiber-reinforced epoxies used in aerospace applications (see *Polymer Matrix Composites with Roughly Equiaxed Reinforcements: Microstructure and Mechanical Behavior**; *Ceramic Matrix Composites with Roughly Equiaxed Reinforcements: Microstructure and Mechanical Behavior**; *Metal Matrix Composites with Roughly Equiaxed Reinforcements: Microstructure and Mechanical Behavior**, *Laminates: Physical and Mechanical Behavior**; *Polymer Matrix Composites: Applications**; *Metal Matrix Composites: Applications**; *Ceramic Matrix Composites: Applications**). Such directionality, which usually results in low strength through the thickness of plate-like constructions, causes particular problems and the emphasis when designing load-carrying joints is on minimizing the stresses in these weak directions.

1. Joint Configurations

1.1 Standard Configurations

The type of joint most often encountered is the lap joint, which has the four basic forms shown in Fig. 1; single lap, double lap, stepped lap, and scarf. Such joints are mainly used between fiber-reinforced plastic (FRP) laminates, but in the stepped and scarf forms could be between an FRP laminate and a metal plate.

Flange joints are also commonly encountered, although there is a shortage of generally available

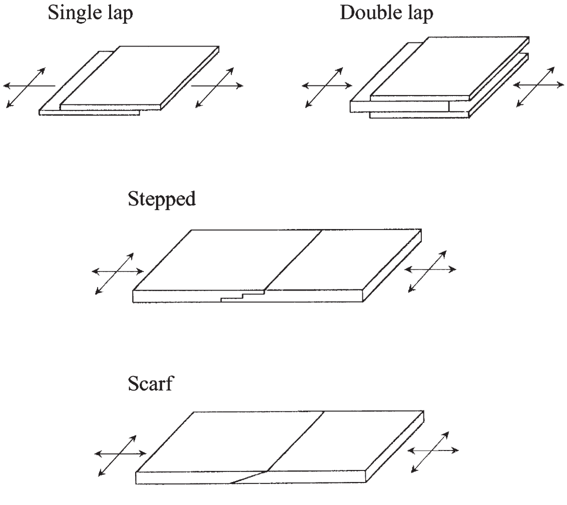


Figure 1
Lap joint configurations for in-plane loading.

information. Such joints would be used to join sections of pipes (tubes), cylinders (e.g., turbojet casings), or the (usually flat) panels that form the sides of containers, for holding water, etc. Tubular components can also be joined by cylindrical forms of the flat lap joints.

1.2 Special Configurations

Nonstandard configurations are reserved for particular applications, e.g., wedge joints are used for securing turbine blades or propeller blades to the turbine disc or propeller hub, respectively. As the rotational speed of the device increases so does the corresponding radial force on the blade and the constraint afforded by the wedge produces an ever tightening joint. The same principle is encountered in the wedge grips of tensile testing machines.

Another special configuration, which has been used for securing helicopter rotor blades to the rotor hub, is the "race track" or "wrapped" joint. Here, continuous fibers are positioned around the hole (Fig. 2) in which the securing pin is located, the forces on the joint being such as to load the fibers in tension, an ideal uni-axial stress state.

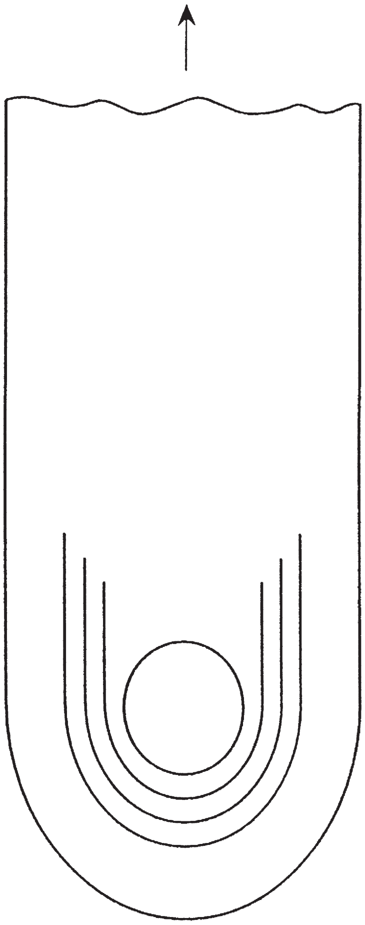


Figure 2
"Wrapped" joint.

2. Load Transfer

The purpose of a structural joint is to transfer load from one side of a joint to the other. In so doing a complex state of stress, which varies throughout the joint, will be set up. The main points can be illustrated by reference to the single cover joint (end-to-end single lap joints) shown in Fig. 3. The load in plate A has to be transferred to plate B via cover plate C. As a consequence the joining medium (adhesive in Fig. 3(b)) will deform in shear, with maximum stresses occurring at the ends of the overlaps. Because the line of action of the force in A and B is not aligned with that in C, overall bending, as shown in Fig. 3(c) will occur. The ends of the overlaps will move apart, setting up through-thickness tensile (so-called, "peel") stress in the joining medium and the plates. It is these stresses mentioned above that are likely to be the cause of failure in FRP laminates.

With flange joints the usual consideration is to maintain a leak-proof configuration by ensuring that the mating faces of the flanges do not separate. Such issues are well understood from many decades of joining steel pipes. With composites, the usual problem is to prevent failure either at the corner of the flange due to the gross bending, or due to the connecting fasteners pulling through the material (Godwin *et al.* 1986).

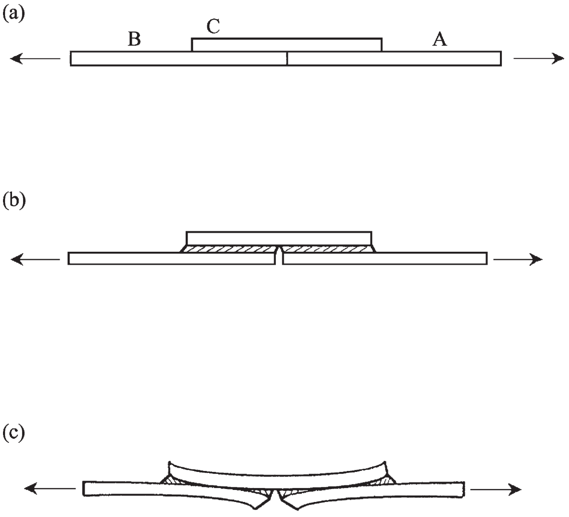


Figure 3
Load transfer across a single cover joint.

3. Joining Polymer Matrix Composites

Polymer matrix composites (PMCs), or FRPs, constitute the bulk of the commercial market (see *Polymer Matrix Composites: Applications**). The majority of FRPs will have a thermosetting matrix and their popularity is due to their ease of processing. These materials are normally used in the form of laminates, with continuous or long fibers, and hence joints are usually of lap configuration. FRPs with thermoplastic matrices are also gaining popularity, although these are more commonly produced with short fibers by processes such as injection molding and hence, rather than being plate-like, will have a bulk form. Joining procedures are then similar to those used for bulk isotropic materials, or for MMCs (see Sect. 4).

Recently introduced materials for aerospace use includes laminates of aluminum alloy sheets bonded to sheets of aramid fiber/epoxy (ARALL) or to sheets of glass fiber/epoxy (GLARE) (Wu *et al.* 1992). Such hybrid laminates would be joined in the same way as conventional FRP laminates.

3.1 Adhesive Bonding

Adhesive bonding (or gluing) is used to join (pre)-cured laminates. A layer of adhesive is inserted between the laminates in the joint region and allowed to cure. So-called "cobonding" is a similar procedure, but here the joint is between uncured laminates, and the adhesive and matrix cure in the same operation. Slightly different, but producing identical configurations, is "cocuring" when there is no additional adhesive layer between parts and the matrix resin acts as the glue during the cure process.

The bonding procedure is perfectly straightforward and well-understood (Matthews 1994). However, care is needed to prepare the mating surfaces by removing all traces of mold release agent, or other contaminant, from the composite. If the composite is to be bonded to a metal then the latter will also need pretreatment to remove grease, etc., and any surface oxide layer (for aluminum and titanium). Adhesives are normally epoxy resin, but could be acrylic, phenolic, or polyurethane. They come in liquid, paste or film form and cure at room temperature, 120 °C or 170 °C.

Of the standard configurations the single lap is the weakest, followed by the double lap, then the stepped lap, with the scarf being the strongest joint (Matthews 1994). Essentially this is due to the eccentricity of load paths being greatest for the single lap and zero for the scarf. Indeed, by making the angle sufficiently small, it is possible to make a scarf joint stronger than the material being joined (efficiency of 100%). The only drawback is the need to maintain a "knife edge" at the corner of the scarf and this is difficult to produce, and avoid being broken off, in wide joints (i.e., items bigger than laboratory test specimens). The stepped joint is difficult and costly to fabricate and so would only be used in the most demanding applications.

The performance of single and double lap joints is enhanced by increasing the length of the overlap region. Further improvements can be brought about by tapering the thickness (being a minimum at the overlap end) of the laminate. Performance is also enhanced by using a ductile adhesive and having the largest possible in-plane stiffness of the laminate, consistent with other structural constraints (Hart-Smith 1972).

Typical applications of bonded FRP joints include: secondary and tertiary (control surfaces, fairings) as well as primary (tail plane-to-fuselage) aircraft structures (Anon 1990, Hoskin and Baker 1986), the chassis of Formula 1 race cars, satellite structures, ship hulls, the doors of truck cabs, suspension bridge decks, and indeed any situation where a smooth profile is desirable and/or drilling holes for fasteners is not feasible (Phillips 1989).

3.2 Mechanical Fastening

Mechanically fastened joints are used in aircraft (e.g., spar-to-skin joints in wings), ships (frame-to-hull joints), and race cars (engine-to-chassis joints). In the aircraft case they are a necessary feature of the overall assembly process (Anon 1990). Mechanical fastening is a suitable procedure when bonding is either difficult (e.g., with thermoplastics) or its durability is questionable. Mechanical fasteners are ideal if disassembly is required, and could be an acceptable approach for (temporary) repair. Of course, fasteners are heavy so a weight penalty will be incurred.

Typical fasteners are screws (only used for the most lightly loaded joints), rivets (limited to laminates no thicker than 6 mm), and bolts (Matthews 1994). The latter produce the strongest joints. Of course, it is necessary to drill holes in the laminate to receive the fasteners, and these holes will give rise to stress concentrations, which inevitably limit the strength of the joint. Drilling the holes requires diamond or carbide tipped bits and controlled feed rates to prevent ply separation on the back face of the laminate.

Mechanical fasteners are mainly used for single lap joints (rivets and bolts), double lap joints (bolts), and, of course, for flanges (bolts). The fiber/matrix combination is very important; carbon (graphite) fiber/epoxy (CFRP) gives the strongest joints, followed by glass fiber/epoxy (GFRP), with "Kevlar" fiber/epoxy (KFRP) being the weakest. A joint with optimum dimensions containing one fully tightened bolt could have an efficiency of around 40% (Hart-Smith 1987). The obvious way to overcome this low performance is to increase the thickness of the laminate in the region of the joint.

Other parameters that can influence performance are the dimensions of the joint, i.e., the fastener diameter, thickness of the laminate, distance of the hole center from the end of the joint, and width of the joint (for one fastener) or the pitch between fasteners for an array of fasteners (Matthews 1994).

3.3 Welding

Because of their low surface energy, thermoplastic matrix composites (TPMCs) are difficult to bond, although with appropriate pretreatment, such as corona discharge, bond strengths are equivalent to those obtained with thermosets (Kinloch *et al.* 1992). The welding process involves melting of the matrix polymer alone and/or an intermediate layer of unreinforced polymer. The general characteristics of welded joints will be similar to those of bonded joints.

The processes that have been considered include: hot plate, vibration, ultrasonic, resistive implant and induction welding (Wise 1992), and, more recently, microwave welding. Induction welding appears to have the most potential for large-scale use and has been employed successfully to fabricate wing components on Airbus airliners (Clements 2000). With all these processes it is possible to weld small specimens, the difficulty is in applying the procedures to real components when large areas have to be heated and consolidated.

4. Joining Metal Matrix Composites

Although not so widely used as PMCs, MMCs have been employed in some key areas, and are becoming more popular (see *Metal Matrix Composites: Applications**). Applications include: automotive drive shafts, valves, pistons, cylinder liners and brake discs; mountain bike frames and golf clubs; struts in space craft; electronic instrument racks; and yacht fittings (Ellis 1994, Koczak *et al.* 1993).

MMCs are normally reinforced by particles or short fibers and hence do not display the anisotropy seen with long fiber reinforcement. They could, clearly, be joined by mechanical fastening or adhesive bonding (provided the latter is compatible with the operation temperature).

However, the main attraction is to be able to use conventional metallurgical techniques such as fusion, forging, and brazing. It is important to ensure that the joining process does not react adversely with the MMC so that properties are reduced in the joint region. The matrix metal is a key determinant of performance, in the latter respect; aluminum (Al) is the most commonly encountered, but titanium (Ti), magnesium (Mg), and other metals have been used (Koczak *et al.* 1993).

4.1 Fusion Welding

During welding the matrix will be above its melting point, but the reinforcement (typically silicon carbide, SiC, or alumina, Al_2O_3) will remain solid. The defects normally found when welding unreinforced alloys, such as solidification cracking, may still be encountered with MMCs, but with the latter additional problems occur, i.e., the higher viscosity of MMC melts; segregation on solidification; reinforcement/matrix interactions; evolution of occluded gases (Ellis 1994).

There are several processes available for fusion welding of MMCs: tungsten inert gas (TIG), metal inert gas (MIG), electron beam (EB), laser beam (LB), resistance, capacitor discharge, plasma (Ellis 1994). TIG and MIG processes should be used when weld fillers are needed. EB and LB methods are attractive because they have short thermal cycles and low heat input.

4.2 Forging and Solid State Welding

Since melting is largely absent, solid state processes avoid the problems encountered with fusion welding of MMCs.

Diffusion bonding is one main category. In this approach the components are held under load at elevated (below melting) temperature in a protective atmosphere or vacuum. Solid phase diffusion bonding (SPDB) is one variant of the process, although transient liquid phase diffusion bonding (TLPDB) is a more successful approach.

Friction welding is another solid state joining process. Friction is created by moving one part against the other, or by a separate tool, using rotary (stir welding), linear, or orbital motion. After sufficient time has elapsed the motion is stopped and the parts brought together under load for the forging stage of the weld. Drive shafts are joined by friction welding. With some MMCs the particulate or fibrous reinforcement close to the bond line may be fragmented.

4.3 Brazing

Since brazing does not involve melting of the matrix, problems of matrix/reinforcement reactions and reinforcement destruction are reduced. Dissimilar materials can be joined by brazing and heat induced distortion is minimized.

5. Joining Ceramic Matrix Composites

CMCs are used in high temperature applications such as turbine blades, piston engine valves, and turbo charger components. It may be necessary to join CMCs to themselves or to a metal (Fernie *et al.* 1991). A particular problem is the brittleness of CMCs, which means that tensile stresses in the joint must be kept to an absolute minimum. The use of mechanical fasteners is normally not feasible, because of the associated stress concentrations, and the difficulty of making holes in such hard materials.

One approach is to use shrink fitting, in which differences in thermal contraction (on cooling from elevated temperature) are exploited to form a tight joint. This method is used in the automotive industry to produce spark plugs (which have an alumina insulator).

Friction welding has also been used to join CMCs to metals, as have SPDB and TLPDB. Brazing can also be used, although this may be difficult because ceramics are not easily wetted by the braze metals. Finally it is possible to join CMCs using ceramic adhesives. However, such materials have very low toughness so joint strengths are generally low.

6. *Concluding Remarks*

All forms of composite can be satisfactorily joined. Adhesive bonding and mechanical fastening are the main methods for PMCs, and these are now mature technologies with well-established design rules.

MMCs can be joined by a variety of methods, but care is needed to avoid unwanted reactions, particularly with fusion processes, and destruction of the reinforcement in the joint region. CMCs are the most difficult type of composite to join and development is needed. Shrink fitting can give an acceptable joint in the right circumstances.

Bibliography

Anon 1990 Case studies. Middleton D H (ed.) *Composite Materials in Aircraft Structures*. Longman, Harlow, UK, Chap. 14

Clements L L 2000 Thermoplastic-matrix components lightens commercial transport. *High Performance Compos*. **8**, 27–31

Ellis M B D 1994 *Joining Metal Matrix Composites—A Review*. Research Report 489/1994. The Welding Institute, Abington, Cambridge

Fernie J A, Threadgill P L, Watson M N 1991 Progress in joining of advanced materials. *Weld. Met. Fabric*. **59**, 179–84

Godwin E W, Matthews F L, Kilty P F 1986 The design of bolted flange joints in GRP. *Plastics Rubber Process. Applic*. **6**, 161–7

Hart-Smith L J 1972 Design and analysis of adhesive-bonded joints. In: *USAF Conf. Proc. AFFDL-TR-72-130*. US Air Force Flight Dynamics Laboratory, Dayton, OH, Paper 6059A

Hart-Smith L J 1987 Design and empirical analysis of bolted or riveted joints. In: Matthews F L (ed.) *Joining Fibre Reinforced Plastics*. Elsevier Applied Science, Barking, UK, Chap. 6

Hoskin B C, Baker A A 1986 Aircraft applications. In: Hoskin B C, Baker A A (eds.) *Composite Materials for Aircraft Structures*. American Institute for Aeronautics and Astronautics, New York, Chap. 13

Kinloch A J, Kodokian G K A, Watts J F 1992 The adhesion of thermoplastic fibre composites. *Philos. Trans. R. Soc. London, Ser. A* **338**, 83–112

Koczak M J, Khatri S C, Allison J E, Bader M G 1993 Metal-matrix composites for ground vehicle, aerospace, and industrial application. In: Suresh S, Mortensen A, Needleman A (eds.) *Fundamentals of Metal Matrix Composites*. Butterworth-Heinemann, MA, Chap. 16

Matthews F L 1994 Bonded and mechanically fastened joints. In: Hollaway L (ed.) *Handbook of Polymer Composites for Engineers*. Woodhead Publishing, Cambridge Chap. 8

Phillips L N 1989 Applications. In: Phillips L N (ed.) *Design With Advanced Composite Materials*. The Design Council, London, Chap. 7

Wise R 1992 Approaching total joining technology for carbon fibre reinforced plastics. *Weld. Inst. Bull*. **1**(January/February), 4–6

Wu H F, Bucci R, Slagter W, Vogelsang B 1992 Bearing strength characteristics of fiber-reinforced aluminum laminates. In: *Proc Composites Testing and Standardisation*. European Association for Composite Materials, Bordeaux, France, pp. 453–470

F. L. Matthews
Imperial College of Science, Technology and Medicine, London, UK

Composites, Microstructure of: Quantitative Description

The inner morphology of materials is generally set at a scale that requires the use of microscopes for observation. Two modes of observation exist in microscopy: transmission microscopy, which requires that the material be transparent, and reflection microscopy, which is often used with composite materials since most are generally quite opaque to light or electron beams. In the vast majority of cases, what is obtained from a microscope is a two-dimensional picture, representing a thin slice (in transmission microscopy) or a planar cut (in reflection) through a three-dimensional structure. Understanding the nature of, and quantifying, the inner microstructure of materials therefore is largely centered around the question of "guessing" and quantifying three-dimensional structures from these two-dimensional images. With composites, such "guessing" is aided by the fact that the shape of the reinforcement is to some extent known; however, although more easily achieved in composites, quantification of microstructures nonetheless remains challenging.

1. *Specimen Preparation and Image Acquisition*

Depending on the scale of the microstructure and the characteristics of the material, optical or scanning electron microscopes will be most often used. Observation by optical microscopy is limited by the optical resolution (0.3 μm for the best lenses). To a lesser extent, observation is also limited in terms of the size of the frame of measurements exploitable from the metrological point of view. Vignetting or geometrical distorsions will appear if the frame is too large.

The scanning electron microscope (SEM) enables work at higher magnification. It possesses several modes, the most classical being back-scattered electron imaging, secondary electron imaging, x-ray elemental mapping, and more recently grain orientation imaging by electron back-channelling analysis. One can obtain many images of the same area using these different observation modes, which can be very useful for the segmentation step described below. Therefore one often works with multimodal images (in optical microscopy a color image is considered as multimodal).

Quantitative and automatic investigation of microstructures requires significant care during specimen preparation. Surface planarity and the absence of scratches or other artefacts potentially introduced during the polishing step are important, as is the choice of any chemical or thermal etching procedure used to enhance certain features of the microstructure. The success of quantitative microstructural analysis is strongly dependent on initial sample preparation, which is often the most important step in the procedure.

2. *Principles of Image Processing and Analysis*

Different methods can be used to treat images. There are methods derived from signal processing (Rosenfeld 1986, Gonzalez and Wintz 1987, Pratt 1991), methods constructed from mathematical morphology analysis (Serra 1982, Coster and Chermant 1985, 1989, Dougherty 1992), and also, to a lesser extent, methods based on pattern recognition and artificial intelligence.

2.1 Image Filtering

Filtering allows elimination of noise, to simplify images, and to facilitate the segmentation procedure (see below). Filters can be classified according to their properties into three categories: linear filters that use the convolution product or Fast Fourier Transform (FFT) (Rosenfeld 1986, Gonzalez and Wintz 1987, Pratt 1991), morphological filters (Serra 1982, Coster and Chermant 1985, 1989) based on the mathematical morphology operators, and rank filters (Kunt *et al.* 1993), which are constructed from the distribution of pixels in a given neighborhood.

Image acquisition systems give images that can be noisy. Generally the noise corresponds to frequencies higher than the signal containing the information. One uses low-pass filters to eliminate such noise; such filters can be of three types: (i) mean linear filters (weighted or not); (ii) opening, closing and their combination (low pass morphological filters); and (iii) median and/or rank filters of the same order.

The contours of objects delineate important transitions in the image; hence, these correspond to high frequency signals. To emphasize these contours, high-pass filters must be used. These come in several types: linear gradient filters (Sobel, Prewitt, etc.), morphological gradient filters (residual of a dilation or an erosion) and Laplacian filters (morphological or linear filter).

2.2 Extraction of Phases or Objects

Most often, on images of microstructures the attribute that allows one set of objects to be distinguished from another is the grey tone level or the color. This is why segmentation by thresholding is generally used to delineate objects in the microstructure. The result of this type of procedure is generally a binary image in which the objects have pixels set equal to 1 while background pixels equal 0. This thresholding operation can be performed either manually or automatically. Other methods can also be used: region growing or contour tracking; these are detailed below.

(*a*) *Manual thresholding*. In manual thresholding, one first chooses an upper and a lower threshold. Objects with a grey tone level within the interval stretching between these thresholds are then extracted from the image. Visual comparison with the original image is used to adjust the threshold levels.

(*b*) *Thresholding from the histogram analysis*. Manual thresholding presents the disadvantage that it depends on the operator, which raises problems of reproducibility and reliability. An automatic thresholding procedure solves these problems, but the choice of the method turns out to be difficult. Nearly all automatic methods are based on histogram analysis (Coster and Chermant 1985, 1989) and can be sorted according to their algorithms, which fall into classes very close to those used in statistical data analysis. One can quote examples such as thresholding by maximization of entropy (Pun 1981, Kapur *et al.* 1985), moment preserving (Tsai 1985), or interclass variance (Otsu 1979).

(*c*) *Thresholding from contours*. In contour thresholding, one uses information derived from object contours. To this end, one has to determine a threshold, which allows detection of a maximum of contours having high contrast and a minimum of contours with low contrast. The first category uses derivative methods (morphological gradient or not, second derivative, Laplacians or optimum filtering: Deriche (1987)). All these methods use thresholding as the last step to eliminate low-gradient transitions. The operators used to detect outlines consist of a threshold adjusted by striking a compromise between the pertinent information and the noise. It is therefore necessary to close incomplete contours afterwards. Three main approaches are used to this end: graph exploration, which searches for the best path between two extremities, the use of neuron networks, and the method of snake contours (Kass and Witkin 1987).

(*d*) *Thresholding from regions.* Two strategies can be used for thresholding from regions: either the growth or the division of regions can be used. A region is defined as a homogeneous zone (defined by texture, grey tone level, color, etc.). Pixel gathering is performed according to a double criterion of adjacence (neighboring at the pixel level) and homogeneity (Lowitz 1983, Chassery and Garbay 1984).

In mathematical morphology the watershed method (Beucher and Meyer 1993) constitutes the main method of segmentation by region growing. The image in grey tone levels is assimilated to a relief made up of crest lines, peaks and valleys. From the minima (markers) of that relief, watersheds are reconstructed. When two catchment basins meet, one obtains the watershed line, which segments the image into several regions. The efficiency of such segmentation is principally linked to the choice of the markers and to the filtering operations involved.

2.3 Binary Image Processing

The binary images obtained after such threshold or segmentation steps can still be modified, again to eliminate noise or artefacts, to separate objects, or to simplify and extract features. Mathematical morphology is the privileged processing tool used to this end. Among the several operators available, the most frequently used are: hole filling, noise elimination by opening–reconstruction, and separation of objects by the watershed technique.

2.4 Last Step: Measurement or Classification

The last step of image processing and analysis consists of either a quantitative description of the microstructure using parameters or functions (vectors) derived from the treated image, or a classification of its different features. To this end, one must take into account the general constraints that limit what one can achieve in quantitative microstructural analysis of materials.

3. *Constraints of Image Analysis in Material Science*

3.1 Statistics

A first general limitation of image analysis is that the area that is sampled is but a small portion of the material from which it is extracted. An important constraint that therefore arises is that of the statistical representativity of the local area that is examined. This factor can, for example, place significant limitations on as simple an image analysis task as the measurement of volume fraction reinforcement in composites.

This issue of statistical representativity of the data gathered is compounded by the fact that to access specific features of the microstructure, a sufficiently high magnification must be used. As all the images are numerically treated, one must in particular be sure that small details represent lines on the pictures that correspond to a minimum of about 20 pixels. In most cases, the choice of the frame size is intimately linked to the magnification. Theoretically one can increase the frame size at constant magnification, using oculars with large optical field, for example. But under these conditions, aberrations mentioned above can become important; hence, one often resorts to fusion of images taken by contiguous views. In the most delicate cases, a compromise between sufficient magnification and statistical representativity must be made.

To guide this compromise, one can use the covariance function method, which tests for data representativity (Serra 1982, Coster and Chermant 1985, 1989). This is easily obtained from mathematical morphological analysis after erosion by a bipoint method as outlined below. The covariance function (Matheron 1965) is calculated as the probability that a random point x and that point $x+h$, obtained by translation of x by a vector $\vec{h}$, both belong to the same phase (single covariance) or to two specific phases (crossed covariance of multiphase materials). When h equals 0, the phase volume fraction, V_V (A), is obtained from the single covariance, and the value obtained for the crossed covariance between the phases A and B, $C(\mathrm{A}, \mathrm{B}, h)$, is zero. In all cases, the slope at the origin informs on the specific surface area of the phase. In the case of a random and stationary structure, the single covariance also tends towards the square of V_V (A) for a large distance h. In the case of crossed covariance, $C(\mathrm{A}, \mathrm{B}, h)$, the asymptote is given by $V_V(\mathrm{A}) \cdot V_V(\mathrm{B})$. If the frame size is not sufficient or if the structure possesses a drift over high length scales, this phenomenon is not observed.

3.2 Geometry

A second limitation in experimental analysis was mentioned in the introduction, namely that microstructural data are generally gathered from two-dimensional sections cut through a three-dimensional material. Unfortunately, there is no general method of knowing or quantifying all features of a three-dimensional microstructure using data gathered on one or a few two-dimensional cuts. To this end, one therefore must make specific hypotheses concerning the nature of the microstructure before using two-dimensional data in constructing the three-dimensional microstructure. In composites, this can be facilitated by prior knowledge of the shape of the reinforcement (fibers, equiaxed particles, spheres, etc.); however, even in these materials this is not a simple question. One method of circumventing this problem is to gather information from a succession of closely spaced planes and build a three-dimensional picture by interpolation (e.g., Li *et al.* 1999). This is, however, an extremely time-intensive procedure; hence, such serial sectioning is very limited in terms of the volume that can be analyzed and, in turn, in terms of the statistical representativity of the data gathered.

When the magnification and the size of the frame of measurements are adequately chosen, one has to try to determine the number of frames to be analyzed to obtain a representative classification or a correct estimate of the measured parameters. Two strategies are possible. First, one supposes that the structure is perfectly stationary even at short dimension scale. In this case, one can determine the minimum number of frames, considering that each parameter measured on a given frame is a random variable that follows a given probability law (normal log, Student law, etc.). Therefore the confidence interval as a function of the number of frames can be measured.

The spatial variability of parameters is not taken into account by such classical statistical methods. With the geostatistic methods (Matheron 1965) the variable is associated with its position: then it is called a regionalized variable and the tool used to analyze its statistical behavior is the variogram, $(\gamma(h) = E(f(x) - f(x+h))^2$, where E is the expectation of the square deviation of the variable measured at x and $(x+h)$. The variogram starts from zero and tends towards an asymptotic value when the frames of analysis become independent. To obtain a variogram it is necessary to implement regularly spaced analysis frames.

4. *Quantification of Microstructures*

Principal criteria of quantification of a microstructure are the phase ratio, the phase size, its state of dispersion, shape, and anisotropy. The relevance of all of these parameters in composites is evident in the examples shown in Fig. 1.

4.1 Stereological Parameters

Stereological parameters (DeHoff and Rhines 1968, Underwood 1970, Coster and Chermant 1985, 1989) are mean magnitude microstructural parameters that correspond to the space in which they are defined. There exists $n+1$ such parameters in n dimensional space. Among these, n stereological parameters correspond to parameters of the lower spaces. Three-dimensional stereological parameters are the volumic fraction of phase X, $V_{\mathrm{V}}(\mathrm{X})$, its specific surface area, $S_{\mathrm{V}}(\mathrm{X})$, the integer of mean curvature per unit volume, $M_{\mathrm{V}}(\mathrm{X})$, and the connectivity number, $N_{\mathrm{V}}(\mathrm{X})$. In two dimensions these are the surface fraction, $A_{\mathrm{A}}(\mathrm{X})$, the specific perimeter, $L_{\mathrm{A}}(\mathrm{X})$, and the connectivity number (number of objects less the number of holes) per unit surface, $N_{\mathrm{A}}(\mathrm{X})$. In a single dimension, these are the linear fraction, $L_{\mathrm{L}}(\mathrm{X})$, and the connectivity number (number of segments) per unit length, $N_{\mathrm{L}}(\mathrm{X})$. In zero dimensional space, there is the point fraction belonging to phase X, $P_{\mathrm{P}}(\mathrm{X})$. These parameters are related by the following stereological relationships:

$$V_{\mathrm{V}}(\mathrm{X}) = A_{\mathrm{A}}(\mathrm{X}) = L_{\mathrm{L}}(\mathrm{X}) = P_{\mathrm{P}}(\mathrm{X})$$

$$S_{\mathrm{V}}(\mathrm{X}) = \pi L_{\mathrm{A}}(\mathrm{X}) = 4N_{\mathrm{L}}(\mathrm{X})$$

$$M_{\mathrm{V}}(\mathrm{X}) = 2\pi N_{\mathrm{A}}(\mathrm{X})$$

These basic stereological parameters allow calculation of derivated parameters such as the mean free path (mean chord), the mean curvature, the mean surface or the contiguity of a phase. The only hypothesis required for validity of all these relations is that the data set must be stationary random.

4.2 Granulometric Measurements

Composite materials generally consist of a large number of isolated reinforcing phase elements such as fibers or particles contained in a continuous matrix. From basic relationships two mean size criteria can be derived to characterize isolated reinforcing phase elements: their mean chord length $\bar{\mathrm{L}}_1 = \mathrm{V_V(X)}/\mathrm{N_L(X)}$ and their mean surface area $\bar{\mathrm{A}}_{\mathrm{A}} = \mathrm{V_V(X)}/\mathrm{N_A(X)}$. These parameters can be completed by granulometric analysis, which leads to the definition of many other possible size parameters. Two strategies are possible to this end. If the phases are granular, individual analysis can be performed by measuring each grain. In that case, many size parameters exist: the diameter of the equivalent circle, the Feret diameters, etc. In the case of an interconnected (matrix) phase, pertinent granulometry may be measured by morphological opening (Serra 1982, Coster and Chermant 1985, 1989). The size parameter is that of the structuring element. It must be noticed that the linear granulometries with the P(l) function obtained from an erosion by a segment is a very powerful tool leading to the measurement of many morphological parameters.

4.3 Phase Dispersion

When a phase is more or less dispersed in a matrix, it becomes of interest to quantify the level of dispersion.

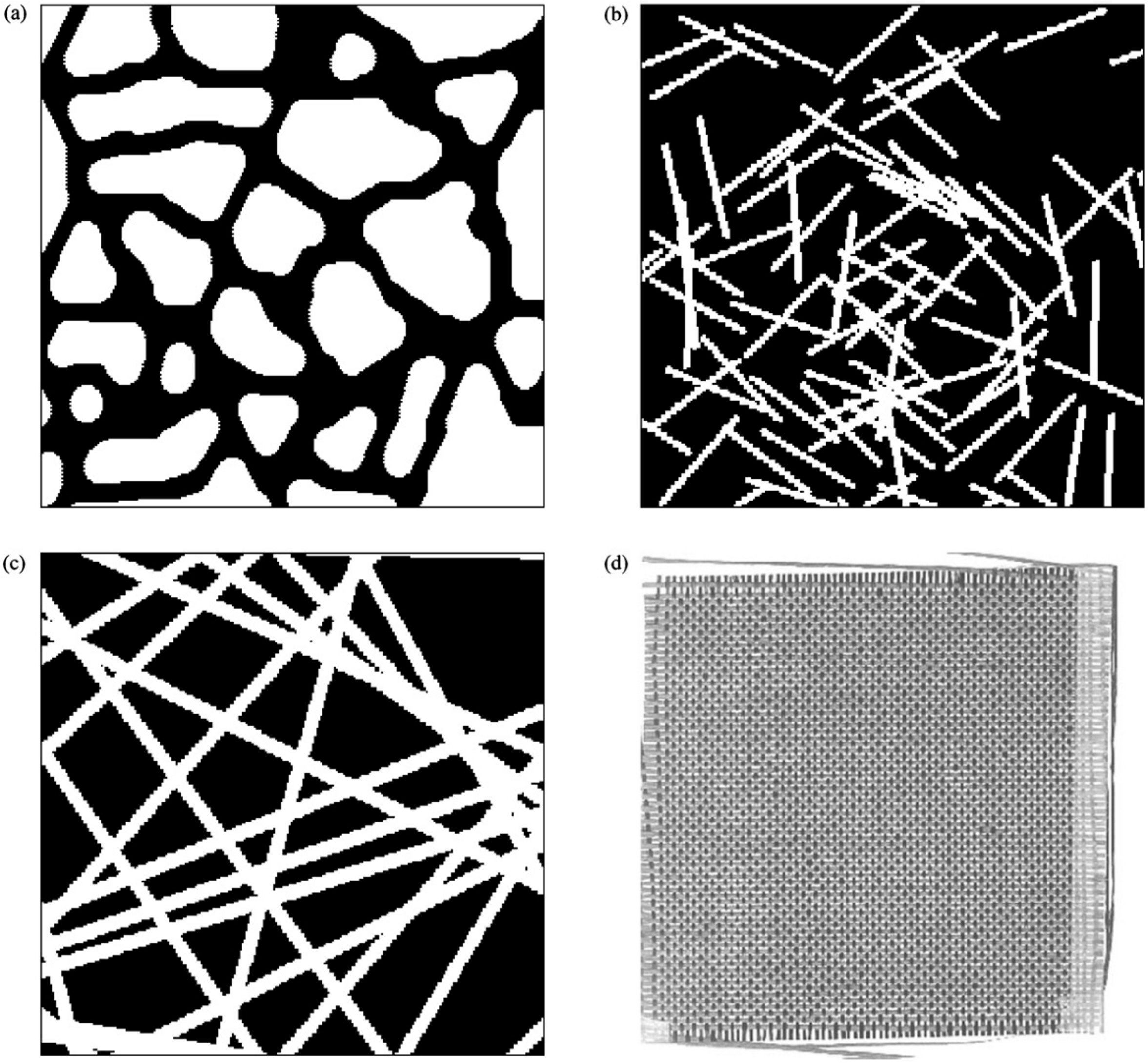


Figure 1
(a) Granular composite; (b) short or chopped fiber; (c) continuous fiber composite; (d) woven fiber composite.

For this purpose one uses the distance function (cartography of the point distances of the matrix from the phase) or the distribution of the nearest neighbors. This covariance function has previously been used to determine if the frame of measurements is large enough. This function can also be used to identify evidence of particle attraction, repulsion, or periodicity, as illustrated in Fig. 2.

4.4 Anisotropy Characterization

To characterize three-dimensional anisotropy, several plane sections (generally orthogonal) must necessarily be analyzed. The choice for the plane orientation depends on *a priori* knowledge of the microstructure. This anisotropy could be described using parameters calculated from stereological quantities defined above (Underwood 1970). The covariance function can also be used to this end; one advantage of this tool is that the orientation of vector $\vec{h}$ can be used to quantify orientation-dependent characteristics of the microstructure.

In polar coordinates, two types of roses can also be plotted to characterize microstructural anisotropy: the rose of directions and of intercepts (Underwood 1970, Coster and Chermant 1985, 1989). The rose of directions is the probability of having a contour

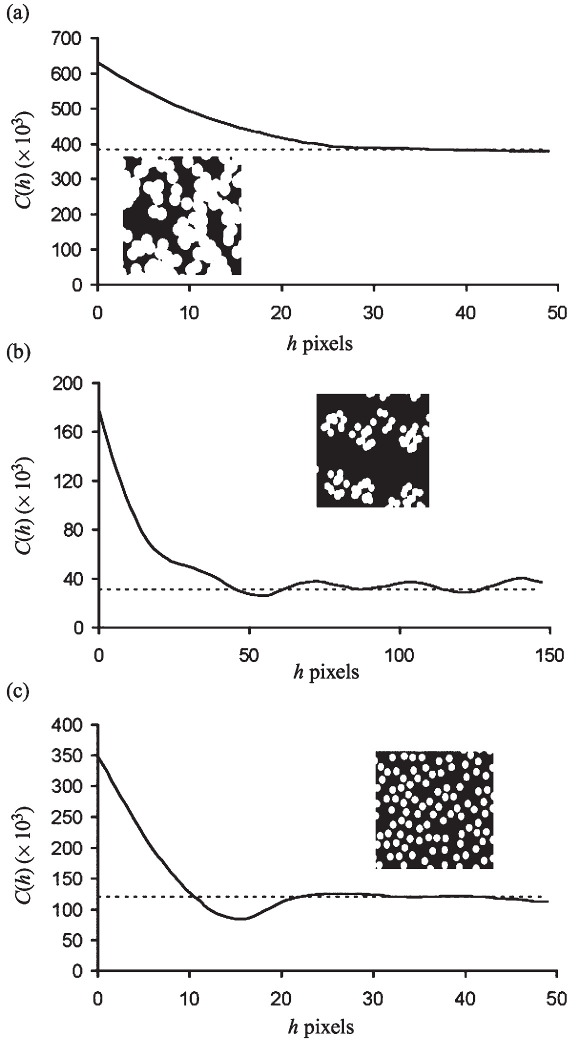


Figure 2
(a) Covariance on a Boolean (random) model; (b) covariance on a cluster (interparticle attraction) model; (c) covariance on a hardcore (interparticle repulsion) model.

oriented along direction α. The rose of intercepts is the projection distribution of the contours in direction α as a function of this angle.

4.5 Shape Analysis

The analysis of shape can only be applied to objects having a well-defined geometry; in composites, this can often be said of the reinforcement. Shape analysis possesses no stereological character, meaning that relevant variables are not simply linked with stereological parameters defined above. For random microstructures, parameters describing the shape are generally dimensionless, since these must be invariant with the magnification. Hence, these are often ratios of two size parameters either measured on an object (for example length/radius for fibers, known as their "aspect ratio"), or ratios of one parameter measured on an object to another measured on its convex envelope (for example surface area of the object/surface area of its convex envelope). Other types of parameters can be used: these can be calculated from the spatial position of pixels belonging to the object (e.g., moments) (Chermant and Coster 1991).

5. *Modeling and Interpretation*

As mentioned above, a full description of a three-dimensional microstructure from two-dimensional cuts is *a priori* impossible. In particular, the number of objects per unit volume, and the distance or the distance distribution of objects constituting dispersed phases in a matrix, cannot be assessed from two-dimensional cuts without making assumptions.

One approach in analysis of composite microstructures is therefore to (i) generate "typical" microstructures of the material at hand, and (ii) to compare this model microstructure with that of the material at hand, somewhat as is practiced in high resolution electron microscopy. If good agreement between model and experimental microstructural data is found, one can draw conclusions concerning the material microstructure using the model.

Due to the more or less random aspect of composite microstructures, probabilistic models appear to be in general well adapted (Serra 1982, Jeulin 1997). Practically all models used to this end are based on a point by point Poisson process. Such probabilistic models can be divided in three categories: space partition, Boolean models and sequential models. The Boolean models are necessarily at least two-phased; the others can be mono- or polyphased.

The best-known space partition model is the Voronoi partition (e.g., Li *et al.* 1999). From a Poisson distribution of points, the influence zone of each point is constructed by drawing lines (in two dimensions) or planes (in three dimensions) equidistant from pairs of neighboring points. A mosaic made of polygons in two dimensions or polyhedra in three dimensions is then obtained. A variant of this model is the Johnson–Mehl model, where points are generated, not via a Poisson process, but by a process comprising nucleation and growth at a constant rate, polygon or polyhedral boundaries being located where growing grains impinge. Unfortunately, such Voronoi partitioning of space does not feature good stereological properties, such that a three-dimensional simulation intersected by a plane must be used to compare the model with two-dimensional microstructural data. Another model of partition is the Poisson mosaic: it is constructed from the intersection of Poisson lines in $\mathbb{R}^2$ or Poisson planes in $\mathbb{R}^3$. In that case the model is stereological.

Boolean models can also be constructed from Poisson points. A primary grain replaces each point and the union of the primary grains corresponds to the Boolean model. The primary grains can take different shapes: monosized spheres, spheres of diameter following a given distribution, polyhedra of known or variable shape, random polyhedra. This model is stereological and three-dimensional parameters can be estimated from plane sections of the material. The parameters are, in that case, the Poisson density and the characteristics of the primary grains (volume, surface area, distribution).

As with Boolean models, sequential models are constructed from a point generation process followed by assignation to each point of grains of chosen shape. However, unlike Boolean models, sequential models are function of time: so the free space and previous grains are covered by new primary grains.

Since composites are made of a reinforcement phase more or less dispersed or agglomerated in a matrix, microstructures will differ according to the reinforcement state of dispersion. With equiaxed reinforcements, several corresponding "granular" model microstructures can be used. The most simple comprises the Boolean scheme issued from a Poisson process, and other schemes incorporating the influence of point to point attraction or repulsion. Figure 1(a) was generated via such a process. Fiber composite microstructures come in various different classes: fibers can be short or continuous, and their orientation can be random or ordered along one, or several directions in woven fiber composites. The case of random fiber structures can be schematically represented by a Boolean scheme of short fibers or by thick Poisson lines for continuous fibers.

The strategy used for analysis depends on the desired characteristics of the model microstructure. In any case, stereological parameters of the material microstructure must be estimated using the methods outlined above, taking into consideration any anisotropy characteristics that may be relevant. Moreover the dispersion state of reinforcing phases should always be analyzed, using the covariance function. Then, these characteristics can be compared with those of model microstructures; for example, covariance plots from a Boolean model show the influence of inter-particle attraction or repulsion on microstructural parameters (Fig. 2). In Fig. 2(a), in the absence of attractive or repulsive effects, a continuous decrease of the covariance towards a constant asymptotic value is observed. In Fig. 2(b), one can see an attractive effect causing the formation of clusters, which translates into oscillations of the covariance around the asymptote at large values of $\boldsymbol{h}$. Finally, a curve extending at large distances below the asymptote provides evidence of interparticle repulsion (Fig. 2(c)), which can be quantified in the model.

Fiber composites can also be subdivided according to the state dispersion of the fibers; in practice, this translates into the difference between random dispersion or woven structures. Random fiber dispersions can be more or less represented by probabilistic models. In the case of short fibers, a Boolean model is generally suitable (Fig. 1(b)). For continuous fiber composites, thick Poisson lines are better adapted (Fig. 1(c)). A drawback of these relatively simple models is that the fibers are superimposed; increased sophistication in how fiber location is computed can of course eliminate this drawback.

With woven composites (Fig. 1(d)), knowledge of the fabrication process gives significant *a priori* information guiding the choice of orientation with regard to the cloth orientation of the plane sections to be considered in analysis. Due to the high anisotropy of such materials, transforming two-dimensional to three-dimensional information can be challenging. It is then better to analyse only sections and not to aim for derivation of three-dimensional stereological parameters other than the volumic ratio. Sections locally perpendicular to yarns can thus be used to gather information concerning fiber distribution or matrix and pore networks. The covariance function can also provide information on the periodicity or nonperiodicity of the fiber architecture.

6. Conclusion

Quantification of microstructures is now a task that can be readily accomplished with the use of computerized analysis of two-dimensional sections through materials. Significant challenges nonetheless exist: sampling must be statistically significant, and the derivation of three-dimensional parameters from two-dimensional pictures is not trivial. Because composite materials in the majority of cases consist of a reinforcement of known geometry dispersed within a continuous matrix, these are particularly amenable to microstructure quantification; general approaches used to this end are outlined in this review.

Bibliography

Beucher S, Meyer F 1993 The morphological approach to segmentation: the watershed transformation. In: Dougherty E R (ed.) *Mathematical Morphology in Image Processing*. Dekker, New York, pp. 433–82

Chassery J M, Garbay C 1984 An iterative segmentation method based on contextual color and shape criterion. *IEEE Trans. PAMI* **6** (6), 794–800

Chermant J L, Coster M 1991 Granulometry and granulomorphy by image analysis. *Acta Stereol.* **10**, 7–23

Coster M, Chermant J L 1985 *Précis d'Analyse d'Images*. Les Editions du CNRS; 2nd edn. 1989. Les Presses du CNRS, Paris

DeHoff R T, Rhines F N 1968 *Quantitative Microscopy*. McGraw-Hill, New York

Deriche R 1987 Using Canny's criteria to derive a recursively implemented optimal edge detector. *Int. J. Comput. Vision* **1** (2), 167–87

Dougherty E 1993 *Mathematical Morphology in Image Processing*. Dekker, New York

Gonzalez R C, Wintz P 1987 *Digital Image Processing*. Addison-Wesley, Reading, MA

Jeulin D 1997 *Advances in Theory and Applications of Random Sets*. World Scientific, Singapore

Kapur J N, Sahoo P K, Wong A K C 1985 A new method for gray-level picture thresholding using the entropy of the histogram. *Comput. Vision Graphics Image Process* **29**, 273–85

Kass M, Witkin A 1987 Analyzing oriented patterns. *Comput. Vision Graphics Image Process* **37**, 362–85

Kunt M, Granlund G, Kocher M 1993 *Traitement Numérique des Images*. Presses Polytechniques et Universitaires Romandes, Lausanne, Switzerland

Li M, Ghosh S, Richmond O, Weiland H, Rouns T N 1999 Three-dimensional characterization and modeling of particle reinforced metal matrix composites: part I. Quantitative description of microstructural morphology. *Mater. Sci. Eng. Ser. A* **265**, 153–73

Lowitz G E 1983 Can a local histogram really map texture information? *Pattern Recognit.* **16**, 141–7

Matheron G 1965 Les Variables Réginalisées et leur Estimation. Masson, Paris

Otsu N 1979 A threshold selection method from gray-level histograms. *IEEE Trans. Syst. Man. Cybern. SMC* **9**, 62–6

Pratt W K 1991 *Digital Image Processing*. Wiley, New York

Pun T 1981 Entropic thresholding, a new approach. *Comput. Vision Graphics Image Process* **16**, 210–39

Rosenfeld A 1986 *Digital Picture Processing*. Academic Press, New York

Serra J 1982 *Image Analysis and Mathematical Morphology*. Academic Press, London

Tsai W H 1985 Moment-preserving thresholding: a new approach. *Comput. Vision Graphics Image Process.* **29**, 377–93

Underwood E E 1970 *Quantitative Stereology*. Addison-Wesley, Reading, MA

J.-L. Chermant and M. Coster
LERMAT, Caen, France

Composites, Physical Properties of

The aim when designing a composite material is to combine the virtues of two or more quite different materials. The extent to which this is achieved depends on the coupling between them. In this article the elastic and physical properties of composites are examined. In this context, predictive models are essential tools because they delineate the properties achievable for a given coupling and thus provide a benchmark against which real composites can be compared. They are also important because they enable the designer to try out different couplings prior to undertaking costly fabrication of new systems.

Models have been especially successful in predicting physical properties (e.g., stiffness, thermal and electrical conductivity, etc.) because, unlike properties such as toughness, they tend to depend on volume average quantities. That toughness is not a volume average quantity is exemplified by the observation that a tough ceramic composite can be made by combining two brittle ceramic materials by careful control of the interfaces. As a result, most physical properties can be modeled fairly accurately on the basis of reinforcement volume fraction, aspect ratio, and orientation data without the need for detailed microstructural information. Throughout this article it is assumed that the microstructure of the composite is fine compared with the size of the test piece or component under consideration. This means that the composite can be treated as homogeneous, i.e., *effective* properties can be considered. Of course this requires that the second phase be distributed in a quasi-uniform manner throughout the composite (either in a regular array or more usually randomly distributed). For the majority of composite systems this is approximately true.

The approaches for modeling thermal conductivity, electrical conductivity, magnetic and electrostatic vector fields, as well as stress tensor fields are all very similar because the governing equations all have a common form (Taya 1990). Each can be written in terms of a flux vector/tensor (Y) and a field vector/tensor (Z) related by a tensor (β) that is material dependent (see also Table 1):

$$\mathrm{Y}_i = \beta_{ij}\mathrm{Z}_j \quad \text{or} \quad \mathrm{Y}_{ij} = \beta_{ijkl}\mathrm{Z}_{kl} \tag{1}$$

According to field theory the flux vector (Y) and field vector (Z) of the four vector quantities conform to:

$$\nabla \cdot \mathrm{Y} = 0 \quad \text{and} \quad \nabla \times \mathrm{Z} = 0 \tag{2}$$

where ∇ is the grad operator ($\mathrm{d}/\mathrm{d}x_1$, $\mathrm{d}/\mathrm{d}x_2$, $\mathrm{d}/\mathrm{d}x_3$) and $\cdot$ and $\times$ denote the inner (dot) and vector (cross) products. These are satisfied for a homogeneous body if Z is written in terms of a scalar potential ϕ ($z = -\nabla\phi$) where ϕ obeys Laplace's equation:

$$\nabla^2\phi = 0 \tag{3}$$

The stress and strain tensors are related by a fourth-rank tensor of materials coefficients (Table 1). In this case $Z_{\mathrm{ij}} = \varepsilon_{\mathrm{ij}} = sym(\nabla u) = \frac{1}{2}(\mathrm{du_i}/\mathrm{dx_j} + \mathrm{du_j}/\mathrm{dx_i})$ where u is the displacement vector. In this case the displacement can be written as $\mathrm{u}_i = \nabla_{\mathrm{i}}(\phi_{\mathrm{o}} + \mathrm{r}_i\phi_j) - 4(1-v)\phi_i$ where v is Poisson's ratio, r_i is the position vector and ϕ_{o} and ϕ_i are scalar and vector potentials that satisfy Laplace's equations (Taya 1990). For inhomogeneous materials, such as composites, $\nabla^2\phi$ becomes nonzero at the interface. The variation of the field variables can be found by using Gauss' theorem to convert from a surface integral at the interface to a volume integral and then solving the

Table 1
Equivalent quantities for the different physical properties.

Physical property	Field vector (tensor), Z	Flux vector (tensor), Y	Material coefficients, β
Stress	Strain tensor, ε_{ij}	Stress tensor, σ_{ij}	Stiffness tensor, S_{ijkl}
Thermal conduction	Temperature gradient, $\nabla_i T$	Heat flux, q_i	Thermal conductivity, Q_{ij}
Electrical conduction	Electric field, E_i	Electric current density, J_i	Electrical conductivity, σ_{ij}
Electrostatics	Electric field, E_i	Dielectric induction, D_I	Dielectric constant, ε_{ij}
Magnetic fields	Magnetic field, H_i	Magnetic flux, B	Magnetic field permeability, μ_{ij}

Source: Taya (1990).

appropriate Green's function subject to certain boundary conditions. Eshelby's analytical solution is applicable to composites containing ellipsoidal inclusions. There is a particularly powerful, elegant, and simple way to evaluate this, which is described in Sect. 4. In most circumstances numerical solutions can be obtained or, alternatively, if detailed information about the composite microstructure is not available then bounds within which composite properties must lie are the best one can obtain.

A framework for describing and understanding many different physical properties is presented here. Rather than take each property in turn, this article is categorized by the different modeling strategies in order to emphasize the areas of commonality between the different properties. The article starts with simple approaches, which can be completed with pencil and paper, moving on to complex ones requiring considerable computing resources. Most assume a perfect coupling between the two or more phases, but in others more complex couplings can be included. In general each section focuses initially on a description of elastic behavior before going on to discuss other physical properties. For each property, the performance of composite systems is examined in the light of the available models and the extent to which they reach their potential is discussed. In the final section a means of identifying promising new composite combinations is presented.

1. Symmetry of Physical Properties

Throughout this article it is assumed that the individual phases are elastically isotropic although in some cases the treatment can be modified to take account of anisotropic matrix or fiber properties (Hashin 1979, Withers 1989). Of course even when the constituents are isotropic, the geometry of the reinforcement and its orientation may confer upon the composite anisotropic properties. Fiber-reinforced composites can exhibit a very wide range of symmetry elements because of the opportunity for different fiber alignments (uniaxial, planar random, laminates, etc. (Table 2)).

Table 2
Symmetry groups representative of different composite reinforcing geometries.

Composite type	Symmetry
Particulate	m4m
Aligned long- or short-fiber composite	4mm
Unidirectional laminate	mmm
Cross-ply laminate or planar random mat or layered material	42m
Angle-ply laminate	222

Source: Newnham (1994).

2. Summation Rules for Composite Properties

The simplest rule for combining the properties of two or more materials in a composite is to assume that the composite obeys some form of summation law of the form:

$$P = f_1 P_1^n + f_2 P_2^n + f_3 P_3^n + \cdots \tag{4}$$

where $f_1, f_2 \ldots$ are the phase volume fractions, P_1 the value of the property for phase one, and the power n describes the nature of the coupling. The most straightforward relation of all is the so-called rule of mixtures (also called the rule of averages) assumption which assumes that the properties of the constituents combine in direct proportion to their volume fraction:

$$P = \sum_i f_i P_i = f_1 P_1 + f_2 P_2 + f_3 P_3 + \cdots \tag{5}$$

2.1 Density

Density (ρ) is one property that almost always obeys the rule of mixtures (ROM) relation:

$$\rho = \sum_i f_i \rho_i = f_1 \rho_1 + f_2 \rho_2 + f_3 \rho_3 + \cdots \tag{6}$$

In many composites the density should be thought of as combining three components, the matrix, the reinforcement, and any porosity. In fact the accurate measurement of density often provides a good means of assessing whether there is any residual porosity after consolidation or porosity introduced by damage or degradation while in service. Although generally true, the ROM relation is not obeyed in cases where a reaction product is formed. Reaction products might be formed deliberately as in reaction-formed ceramic composites; for example, Refil in which silicon is reacted with carbon powder to consolidate around SiC particles. Alternatively, reaction products might be formed accidentally as in the case of metal matrix/ceramic reinforcement reactions (e.g., Ti/SiC composites).

2.2 *Stiffness*

As with many other physical properties, the ROM relation is successful at predicting the axial stiffness response of long-fiber composites. This is because the axial strains (ε_1, ε_2, …) are the same in each phase (the so-called Voigt bound) so that the phase stresses are proportional to their stiffnesses (E_i), and the force (F_i) borne by each phase is $f_i A E_i \varepsilon$ giving:

$$E_{\mathrm{comp}} = (F_1 + F_2 + \cdots)/A\varepsilon = f_1 E_1 + f_2 E_2 + f_3 E_3 + \cdots \tag{7}$$

However, for most phase geometries the strains are not equal and the ROM tends to overestimate the stiffness (Fig. 1). This is because the strains in the more compliant phases tend to be larger than those in the stiffer phases. At the other extreme, the series mixing or Reuss relation (i.e., $n = -1$ in Eqn. (1)) is valid for the normal stiffness of layered systems and is an acceptable approximation for the transverse stiffness of uniaxial fiber composites:

$$E_{\mathrm{trans}}^{-1} = f_1 E_1^{-1} + f_2 E_2^{-1} + f_3 E_3^{-1} + \cdots \tag{8}$$

In the majority of cases, however, the reinforcement shape/arrangement gives rise to more complex couplings, which are neither solely in parallel nor in series. As a result the Voigt and Reuss relations act as extreme bounds (Fig. 1).

2.3 *Thermal Conductivity*

There are many applications for which the high thermal stability and low thermal conductivity coupled with the reasonable thermal shock resistance of ceramic matrix composites makes them attractive. As mentioned in the introduction, the basic physical equations governing thermal conductivity (Q) are rather similar to those for elasticity (Table 1) and for this reason it is possible to apply many of the models

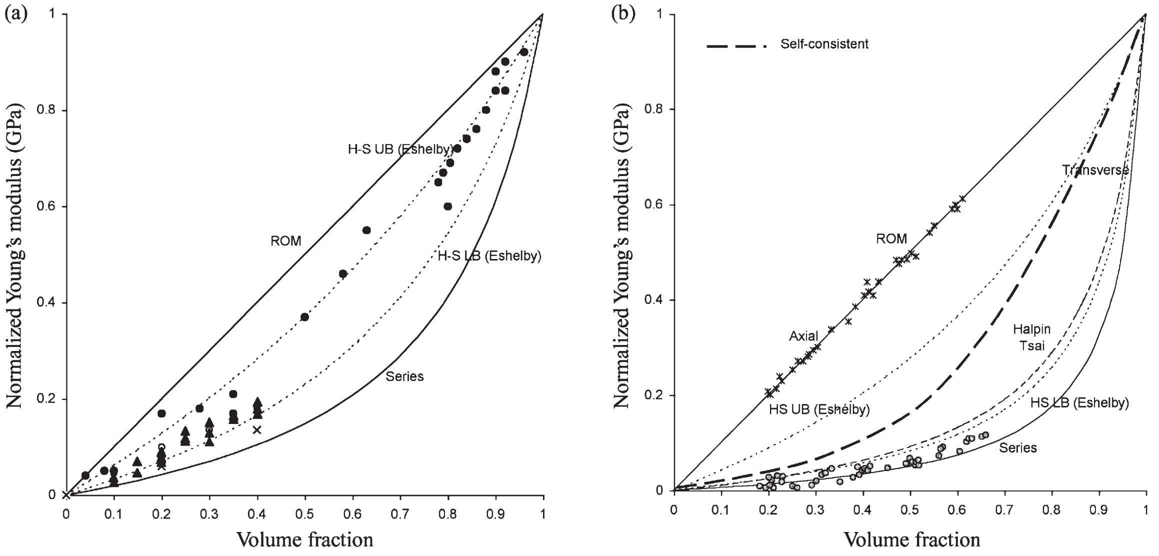


Figure 1
Normalized stiffness ($(E - E_{\mathrm{R}})/(E_{\mathrm{R}} - E_{\mathrm{M}})$) plots including the ROM, series, Halpin–Tsai, Eshelby, self-consistent predictions, and HS bounds (dashed lines) for stiffness ratios characteristic of (a) metal and ceramic particle composites ($E_{\mathrm{R}}/E_{\mathrm{M}} \sim 6$) and (b) polyester/glass-fiber composites ($E_{\mathrm{R}}/E_{\mathrm{M}} \sim 15$). In (a) data is for WC/Co cermet (circles) (Shaw and Uhlmann (1987), alumina/spinel (squares) (Hunter and Bronwell 1967), glass/alumina (crosses) (Davidge and Green 1968), and Al/SiC particles (triangles) (Clyne and Withers 1993). In (b) data is for polyester/glass fiber (Bintrup 1975).

derived for predicting elastic behavior to thermal conductivity.

The conductivity of a composite is usually intermediate between the properties of the two phases. However, only in special cases does the ROM summation rule correctly describe the variation. This is because the heat tends to flow predominantly through the microstructural path of highest conductivity rather than sample each phase equally. Parallel (ROM) and series summation ($n=1$ and -1 in Eqn. (4) respectively) do apply, however, for layered composites parallel and perpendicular to the layers:

$$\begin{aligned} Q_{\text{parallel}} &= f_1 Q_1 + f_2 Q_2 + \cdots \\ 1/Q_{\text{perpendicular}} &= f_1/Q_1 + f_2/Q_2 + \cdots \end{aligned} \tag{9}$$

The ROM also predicts the axial conductivity of fiber composites with good reliability.

3. *Semiempirical Methods*

A number of semiempirical methods have been developed. These combine ease of computation with basic elements of micromechanics, but lack a rigorous fundamental basis.

3.1 Shear Lag Model for Axial Stiffness

The shear lag model is perhaps the simplest model for axial stiffness, which takes into account the geometry of the phases as well as their volume fraction (Cox 1952). It assumes that the load is transferred from the matrix to the fibers by the generation of shear stresses at the fiber/matrix interface. The model predicts an axial composite stiffness given by:

$$E_{\text{axial}} = f_{\text{R}} E_{\text{M}} \left(1 - \frac{\tanh(ns)}{ns}\right) + (1 - f_{\text{R}}) E_{\text{M}} \tag{10}$$

where

$$n = \sqrt{\left(\frac{2E_{\text{M}}}{\text{E}_{\text{R}}(1 - f_{\text{R}})\ln(1/f_{\text{R}})}\right)}$$

Because the model neglects the transfer of stress normal to the fiber ends, the model tends to underestimate the stiffness of very short fiber- and particle-containing composites (Fig. 2), for which the transfer of normal stresses is as important as shear.

3.2 Halpin–Tsai Equation for Transverse Stiffness

The Halpin–Tsai equation provides a means of estimating the elastic modulus, E_{22}, of a long-fiber composite transverse to the fiber alignment direction:

$$E_{22} = \frac{E_{\text{M}}(1 + \xi \eta f_{\text{R}})}{(1 - \eta f_{\text{R}})} \tag{11}$$

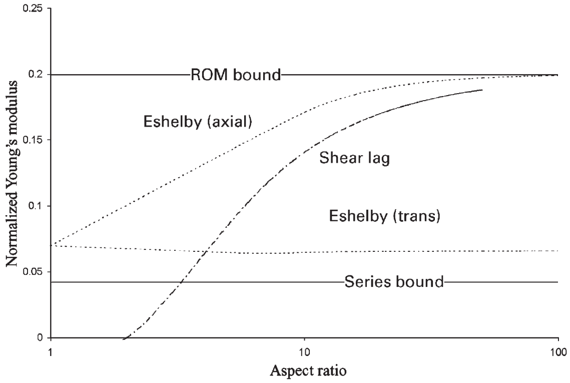


Figure 2
The variation in the predicted axial and transverse stiffness as a function of particle aspect ratio for a 20% alumina (400 GPa)-containing glass (75 GPa) matrix composite. This illustrates how the Eshelby and shear lag models predict an increase in axial stiffness with increasing aspect ratio. Note that both the shear lag and the Eshelby models tend to the ROM estimate at high ratios and that the series (Reuss) bound provides an underestimate of transverse stiffness at all aspect ratios.

where

$$\eta = \frac{E_{\text{R}} - E_{\text{M}}}{E_{\text{R}} + \xi E_{\text{M}}}$$

and ξ is an adjustable parameter, dependent on the fiber arrangement, of order unity. Although it is not founded on rigorous elasticity theory, it does express the increased load bearing of the reinforcing phase relative to the equal stress relation. It is correct at the extremes of reinforcement content and is usually found to be a good estimate over the whole range of reinforcement (Hull and Clyne 1996).

3.3 Stiffness Relations for Porous Materials

While general purpose models for composite stiffness can be applied to porous materials simply by setting the stiffness of the reinforcing phase to zero, a number of models have been developed specifically to describe porous materials, such as the empirical form:

$$E = E_{\text{M}}(1 - af_{\text{p}} + bf_{\text{p}}^2) \tag{12}$$

where f_{p} is the pore fraction and a and b are constants (1.9 and 0.9 respectively for spherical pores (MacKenzie 1950)). Note that at low volume fractions the effect of the pores is essentially twice that predicted by the ROM estimate (i.e., $a=1$, $b=0$) and is in good agreement with the more complex Eshelby relation (Fig. 3). Another model proposed to describe the elastic behavior of concrete is simply $E = E_{\text{M}}(1-f_{\text{p}})^3$ (Weidmann *et al.* 1990), for which the pores are three

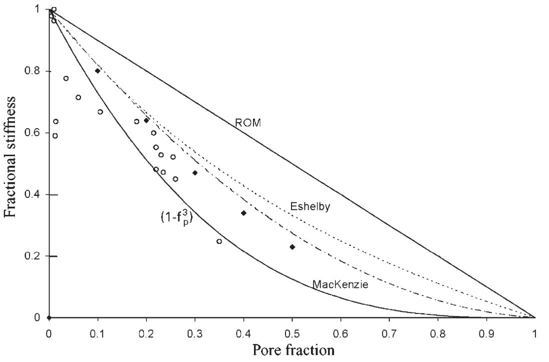


Figure 3
Stiffness of porous ceramics (diamonds, Kingery *et al.* (1976); circles, Larsen (1979)) as compared with the Eshelby, ROM, MacKenzie, and $(11-f_p^3)$ models.

times as important as that predicted by ROM at low pore fractions.

3.4 *Thermal Microstresses*

In two-phase materials the propensity for large thermal microstresses is considerable and it is important to be able to estimate their likely magnitude. For spherical inclusions, the pressure P (or average reinforcement stress $<\sigma>_R$) generated within the inclusions by a matrix/inclusion misfit $\Delta\alpha\Delta T$ is given y:

$$P=\frac{4G_M\dfrac{K_R(1+\nu_M)}{3K_M(1-\nu_M)}}{1+\dfrac{(K_R-K_M)(1+\nu_M)}{3K_M(1-\nu_M)}}\Delta\alpha\Delta T=-<\sigma>_R \qquad (13)$$

Simpler alternative approaches include $P=\Delta\alpha\Delta T/((1+\nu_M)/2E_M+(1+\nu_p)/2E_p)$ (Selsing 1961) designed for isolated crystals in glass matrices, and $\sigma_i=K(\alpha-\alpha_i)\Delta T$ where K and α are the composite bulk modulus and thermal expansion coefficients and α_i the expansion coefficient of the phase under consideration.

4. *The Eshelby Model*

Perhaps the most important rigorous analytical model for the distribution of stress (and many other properties) in a composite is the Eshelby model. Strictly, it applies only to the placement of a single ellipsoidal inclusion in an infinite matrix. The realization that the method could be extended to composites containing practically useful reinforcement volume fractions was made by Brown and Stobbs (1971) and Mori and Tanaka (1973). Using their mean field approach in which an inclusion experiences the average effect of the neighboring inclusions means that it is not a strict geometrical model. But since the distribution of the phases is usually chaotic and impractical to determine anyway, a model which does not require accurate reinforcement coordinates has distinct advantages. The model can make accurate predictions for oblate and prolate spheroids at volume fractions below 30% beyond which self-consistent approaches are required (Sect. 6).

4.1 *Stiffness*

The stiffness tensor, **C**, for the composite is given by:

$$\mathbf{C}=\left(\mathbf{C}_M^{-1}-f_R\{(\mathbf{C}_R-\mathbf{C}_M)[\mathbf{S}-f_R(\mathbf{S}-\mathbf{I})]+\mathbf{C}_M\}^{-1}\times(\mathbf{C}_R-\mathbf{C}_M)\mathbf{C}_M^{-1}\right)^{-1}. \qquad (14)$$

where **I** is the identity matrix,$\mathbf{C}_M$ and $\mathbf{C}_R$ the stiffness tensors for matrix and reinforcement, and **S** is the Eshelby tensor which is given in Clyne and Withers (1993) along with computer programs for stiffness and internal stress calculation. Its predictive capability is better than shear lag models and the errors associated with applying it to nonellipsoidal geometries (cylinders, plates, etc.) are small.

For axial loading of a spherical-particle-containing system, Eqn. (14) leads to (Clyne and Withers 1993):

$$E=\frac{E_M}{1+f_R[(K_M-K_R)/3K_R8G_M+(G_M-G_R)/3G_RG_M]} \qquad (15)$$

where G and K are the shear and bulk modulus, respectively. Eshelby's approach can also be used to model the stiffness of systems in which the interface has no strength, the particles crack, or the particles crumble. In this context, the model has been used to assess the development of damage in composite systems as a function of loading by the progressive fall-off in the measured stiffness (Brechet *et al.* 1991).

4.2 *Thermal Conductivity*

The composite conductivity tensor is given by:

$$\mathbf{Q}=\left(\mathbf{Q}_M^{-1}-f_R\{(\mathbf{Q}_R-\mathbf{Q}_M)[\mathbf{S}-f_R(\mathbf{S}-\mathbf{I})]+\mathbf{Q}_M\}^{-1}\times(\mathbf{Q}_R-\mathbf{Q}_M)\mathbf{Q}_M^{-1}\right) \qquad (16)$$

which is directly analogous to Eqn. (14), but with the **S** tensor now appropriate for conductivity (Clyne and Withers 1993) and $\mathbf{Q}_R$ and $\mathbf{Q}_M$ the reinforcement and matrix conductivity tensors. The variation with reinforcement loading is shown in Fig. 4.

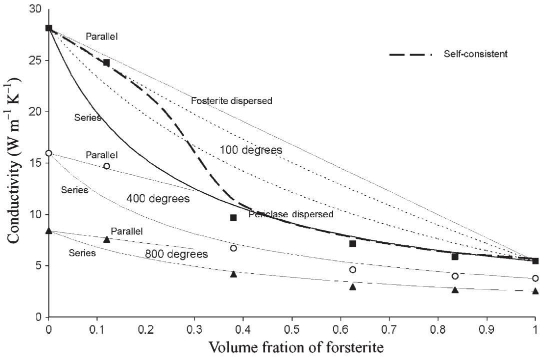


Figure 4
Thermal conductivities of periclase (MgO)/fosterite (Mg_2SiO_4) ceramic composites of different fosterite volume fractions at 100 °C, 400 °C, and 800 °C compared with various model predictions (Kingery *et al.* (1976). In each case the ROM and series predictions are shown with solid lines, the HS bound (Eshelby) predictions with dashed lines.

4.3 *Thermal Expansion and Thermal Microstresses*

Stress balance requires that the volume-averaged mean field thermal stresses $<\sigma_i>$ in the two phases conform to (Fig. 5):

$$(1-f_{\mathrm{R}})<\sigma>_{\mathrm{M}}+f_{\mathrm{R}}<\sigma>_{\mathrm{R}}=0 \tag{17}$$

The stress-free misfit ($\Delta\varepsilon=\Delta\alpha\Delta T$) between the phases upon changing the temperature T is equivalent to Eshelby's stress-free transformation (Clyne and Withers 1993). The matrix stress is then given by:

$$\begin{aligned}<\sigma>_{\mathrm{R}}=&f_{\mathrm{R}}\mathbf{C}_{\mathrm{M}}(\mathbf{S}-\mathbf{I})\{(\mathbf{C}_{\mathrm{M}}-\mathbf{C}_{\mathrm{R}})[\mathbf{S}-f_{\mathrm{R}}(\mathbf{S}-\mathbf{I})]\\&-\mathbf{C}_{\mathrm{M}}\}^{-1}\times\mathbf{C}_{\mathrm{R}}\Delta\alpha\Delta T\end{aligned} \tag{18}$$

Of course this approach assumes that the thermal expansion misfit is accommodated wholly elastically. In many composites there exist many mechanisms by which the thermal stresses can partially relax (creep, plasticity, cracking, etc.). In such cases the equivalent stress-free temperature T_{esf} (Arsenault and Taya 1987) is often useful. This is the temperature that would generate the stresses actually observed were the misfit accommodated completely elastically. A large difference between the actual temperature change and T_{esf} is indicative of active relaxation micromechanisms.

The Eshelby model also provides a good prediction of thermal expansivity:

$$\begin{aligned}\alpha_{(\mathrm{comp}}=&\alpha_{\mathrm{M}}-f_R\{(\mathbf{C}_{\mathrm{M}}-\mathbf{C}_{\mathrm{R}})[\mathbf{S}-f_{\mathrm{R}}(\mathbf{S}-\mathbf{I})]\\&-\mathbf{C}_{\mathrm{M}}\}^{-1}\mathbf{C}_{\mathrm{R}}(\alpha_{\mathrm{R}}-\alpha_{\mathrm{M}})\end{aligned} \tag{19}$$

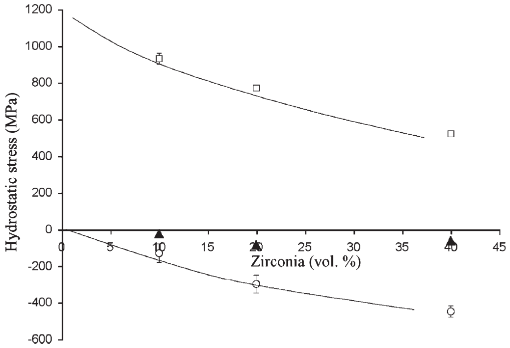


Figure 5
The variation in hydrostatic stresses in Al_2O_3/ZrO_2 composites as a function of the ZrO_2 volume fraction (Wang *et al.* 1984). Triangles delineate the net macrostress and indicate that the thermal stresses are self-equilibrating in agreement with Eqn. (17).

5. *Bounds*

If one is to predict accurately a property of a composite then one needs to know the arrangement/distribution of the phases as well as their volume fractions. Without this one can only place bounds in the range within which that property must lie.

5.1 *Stiffness*

The simplest bounds are the Reuss and Voigt bounds (Sect. 2). These bounds are unnecessarily wide apart for most composite systems and more useful bounds have been developed by Hashin and Shtrikman (1963) among others. Their lower bound is equivalent to the Eshelby model and the upper bound is essentially equivalent to the Eshelby model with the phases reversed, i.e., the contiguous phase having the properties of the reinforcement and the isolated phase those of the matrix.

The upper E_{Uaxial} and lower E_{Laxial} bounds for the axial elastic modulus of a long-fiber composite containing a reinforcement fraction f_{R} are given by (Hill 1964):

$$\begin{aligned}E_{\mathrm{Uaxial}}=&f_{\mathrm{R}}E_{\mathrm{R}}+(1-f_{\mathrm{R}})E_{\mathrm{M}}\\&+\frac{4f_{\mathrm{R}}(1-f_{\mathrm{R}})(\nu_{\mathrm{R}}-\nu_{\mathrm{M}})^2}{f_{\mathrm{R}}/K_{\mathrm{M}}+(1-f_{\mathrm{R}})/K_{\mathrm{R}}+1/G_{\mathrm{R}}}\end{aligned} \tag{20}$$

$$\begin{aligned}E_{\mathrm{Laxial}}=&f_{\mathrm{R}}E_{\mathrm{R}}+(1-f_{\mathrm{R}})E_{\mathrm{M}}\\&+\frac{4f_{\mathrm{R}}(1-f_{\mathrm{R}})(\nu_{\mathrm{R}}-\nu_{\mathrm{M}})^2}{f_{\mathrm{R}}/K_{\mathrm{M}}+(1-f_{\mathrm{R}})/K_{\mathrm{R}}+1/G_{\mathrm{M}}}\end{aligned}$$

where ν_{R} and ν_{M} are Poisson ratios, K_{R} and K_{M} are bulk modulus, and G_{R} and G_{M} are shear modulus of reinforcement and matrix respectively. Note how these bounds collapse to give the ROM result when the Poisson's ratios of the two materials are equal.

Hashin and Shtrikman (HS) (1963) have derived best possible bounds for all the elastic constants (K, G, E, and ν) for an isotropic composite material. The lower bounds for K and G are given by:

$$K_L = K_M + \frac{f_R}{1/(K_R - K_M) + 3(1 - f_R)}(3K_M + 4G_M)$$

$$G_L = G_M + \frac{f_R}{1/(G_R - G_M) + \dfrac{6(K_M + 2G_M)(1 - f_R)}{[5G_M(3K_M + 4G_M)]}} \qquad (21)$$

The corresponding upper bounds can be obtained by reversing the properties of the matrix and reinforcement in the above. Since only two elastic constants are independent, E_L, E_U, ν_L, and ν_U can be determined from Eqn. (21). Such bounds are shown in Fig. 1.

5.2 Thermal Conductivity

For long-fiber composites, the axial conductivity is well described by the ROM (Eqn. (9)). The transverse conductivity lies between the upper Q_{Utrans} and lower Q_{Ltrans} bounds:

$$Q_{\mathrm{Utrans}} = Q_R + \frac{(1 + f_R)Q_R(Q_M - Q_R)}{Q_R + f_R(Q_M - Q_R)/2} \quad \text{and}$$

$$Q_{\mathrm{Ltrans}} = Q_M + \frac{f_R Q_M(Q_R - Q_M)}{Q_M + (1 - f_R)(Q_R - Q_M)/2} \qquad (22)$$

These are analogous to the HS bounds for elasticity and, as before, the lower bound is identical to the Eshelby prediction and the upper bound can be obtained by swapping the properties of the dispersed and contiguous phases. Note that in the limit of $Q_R = 0$ the transverse thermal conductivity is given by $Q_{\mathrm{trans}} = Q_M(1 - f_R)/(1 + f_R)$. By comparison the composite cylinder assembly model for aligned long fibers makes the prediction:

$$Q_{\mathrm{trans}} = Q_M\left(\frac{Q_M(1 - f_R) + Q_R(1 + f_R)}{Q_M(1 + f_R) + Q_R(1 - f_R)}\right) \qquad (23)$$

A relation has been proposed to describe the conductivity of particulate systems (Kingery *et al.* 1976), which is equal to the HS bound:

$$Q = Q_M\left(\frac{1 + 2f_R(Q_R - Q_M)/(2Q_M + Q_R)}{1 - f_R(Q_R - Q_M)/(2Q_M + Q_R)}\right) \qquad (24)$$

5.3 Thermal Expansivity

The relation of Levin (1967):

$$\alpha = \alpha_{\mathrm{ROM}} + \frac{\alpha_M - \alpha_R}{(1/K_M) - (1/K_R)}(1/K - 1/K_{\mathrm{ROM}}) \qquad (25)$$

where the subscript ROM denotes rules of mixtures volume averages, can be combined with HS bounds for bulk modulus to obtain the best possible bounds for the expansion coefficient for isotropic composites (e.g., particulate or random fibers):

$$\alpha_U = \alpha_R - \frac{(\alpha_R - \alpha_M)K_M(1 - f_R)(3K_R + 4G_R)}{K_R(3K_M + 4G_R) + 4(K_M - K_R)G_R(1 - f_R)} \qquad (26)$$

This is identical to the value obtained by the composite sphere assemblage model. Once again the lower bound is obtained by swapping the properties of the phases over.

Turner (1946) derived the following expression:

$$\alpha = \frac{\sum \alpha_i K_i f_i}{\sum K_i f_i} \qquad (27)$$

which assumes that the thermal stresses are not significant. Although not a formal bound, this formulation has been shown to be a reasonable approximation for mullite/coordierite (see Fig. 6). In brittle matrix or reinforcements systems the thermal stresses may cause microcracking to occur. This can give rise to nonlinearities in the thermal expansion curve as well as a hysteresis between the expansion coefficient observed on heating and cooling (Fig. 7).

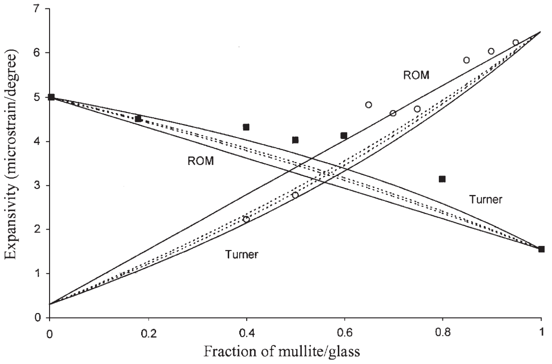


Figure 6
Expansivity as a fraction of silica/glass (circles) (Tummala and Firedber 1970) and ($Al_6Si_2O_{13}$)/cordierite ($Mg_2Al_4Si_5O_{18}$) composite (squares) (Mussler and Shafer 1984). The relatively small difference in elastic properties means that the ROM and Turner models are relatively close together, but note how both lie outside the Levin/HS/Eshelby bounds (dashed lines) (Eqn. (25)).

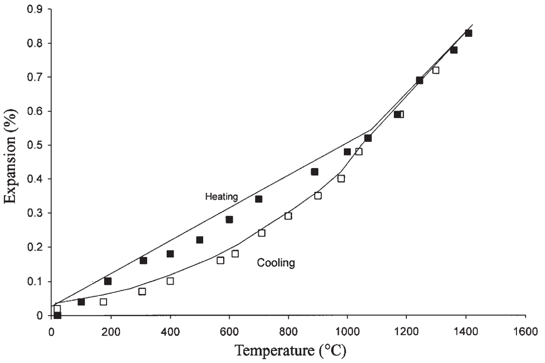


Figure 7
Thermally induced intergranular stresses can cause microcracks leading to hysteresis in the thermal expansion curve, as shown here for TiO_2 (Kingery *et al.* 1976).

For transversely isotropic composites (e.g., aligned fiber composites), Levin (1967) derived the following for the axial and transverse expansion coefficients:

$$\alpha_{\mathrm{axial}} = \alpha_R + (\alpha_M - \alpha_R)[(1+\nu_M)(E_{\mathrm{axial}} - f_R E_R) - (1+\nu_R)(1-f_R)E_M]/[E_{\mathrm{axial}}(\nu_M - \nu_R)] \qquad (28)$$

$$\alpha_{\mathrm{trans}} = (\alpha_M - \alpha_R)[(1+\nu_M)(f_R \nu_{\mathrm{axial}} E_R - \nu_R E_{\mathrm{axial}}) + (1-f_R)\nu_{\mathrm{axial}} E_M (1+\nu_R)]/[E_{\mathrm{axial}}(\nu_M - \nu_R)]$$

These can be used to obtain bounds simply by using the appropriate upper or lower bounds on the composite modulus. Schapery (1968) and Rosen and Hashin (1970) have derived upper and lower bounds for thermal expansivity along similar lines. An equivalent representation is given by:

$$\alpha_{\mathrm{axial}} = \alpha_{\mathrm{ROM}} + \frac{\alpha_R - \alpha_M}{1/K_R - 1/K_M} \times \left(\frac{3(1-2\nu_{\mathrm{axial}})}{E_{\mathrm{axial}}} - 1/K_{\mathrm{series}} \right) \qquad (29)$$

$$\alpha_{\mathrm{trans}} = \alpha + \frac{\alpha_R - \alpha_M}{1/K_R - 1/K_M} \times \left(3/2K_{\mathrm{trans}} - \frac{3\nu_{\mathrm{axial}}(1-2\nu_{\mathrm{axial}})}{E_{\mathrm{axial}}} - 1/K_{\mathrm{series}} \right)$$

where $\alpha_{\mathrm{ROM}} = f\alpha_R + (1-f)\alpha_M$ and $1/K_{\mathrm{series}} = f/K_R + (1-f)/K_M$. The lower bound is equivalent to the Eshelby model (Pedersen 1983, Clyne and Withers 1993), while the upper bound is equivalent to the Eshelby model with the phases reversed.

6. *Self-consistent Models*

In Sect. 5 we saw that the Eshelby treatment represented the lower bound for stiffness. However, problems arise at high volume fractions (>30%) because the lower bound is structurally biased. It assumes a connected matrix phase and isolated reinforcements. This picture becomes increasingly unrealistic at high reinforcement loadings when the percolation threshold is breached (between $f = 0.15$ and 0.4) and the reinforcements begin to touch. In simple terms, rather than consider a reinforcing inclusion in a matrix, self-consistent methods use iterative procedures to place the phases in a matrix having the properties of the composite. At low volume fractions self-consistent methods approach the lower (Eshelby) bound, while at high reinforcement fractions they approach the upper bound (Fig. 1). Much the same behavior is observed in Fig. 4 where the actual thermal conductivity of the periclase/fosterite composite at a given temperature transfers from the upper curve (fosterite is the dispersed phase) to the lower (periclase is the dispersed phase) as the volume fraction of periclase is decreased. Such a transition is analogous to that predicted by self-consistent models (e.g., Pedersen and Withers 1992). Furthermore, this is to be expected given the probable contiguity of both phases in the intermediate volume fractions.

7. *Finite Element Modeling*

Finite element modeling is becoming increasingly popular. This is due in part to the increased freedom one has in terms of reinforcement geometry and distribution, interface, and mechanical properties, and in part due to advances in computational power and user-friendly codes. With modern computers elastic models containing many thousands of elements can be constructed and solved quickly. The basic tenet underlying this approach is that one can describe the behavior of the composite through the performance of a representative unit cell. This is necessary because, except for a few idealized cases, the distribution of the phases is ill defined and chaotic. It is not normally possible to determine the complete arrangement of the phases in three dimensions over a volume comparable with mechanical test dimensions, and even if it were, it would not be computationally feasible to build a model with such a level of detail. As a result it is both computationally and experimentally desirable to exploit a unit cell. This may comprise many fibers in a matrix, but it is more usual to use a cell that may only comprise a part of one reinforcement and to rely upon symmetry operations to build up the complete composite structure. Of course the results are dependent on the precise reinforcement shape and distribution selected. On the whole, because thermoelastic properties are volume-averaged linear elastic properties, they are very well determined. In fact thermoelastic properties are relatively insensitive to local details (peak stresses and locations) and for this reason the increased computational complexity involved in using finite element models is not usually warranted from the viewpoint of predicting thermoelastic properties. Of course this is not the case for nonlinear properties such as fatigue and failure.

8. Dielectric Constant and Electrical Conductivity

By convention the symbols for composite dielectric constant (ε) and electrical conductivity (σ) are the same as those used for strain and stress respectively (Table 1) and so the two are treated separately from the mechanically coupled properties here simply to avoid confusion. Because many of the mathematical expressions for conductivity can be obtained simply by exchanging σ for ε in the corresponding equations for dielectric constants, this section is focused primarily on dielectric properties without loss of generality. All of the modeling strategies introduced in the preceding sections can be applied to the prediction of dielectric constants.

At the simplest level the summation rules described by Eqn. (4) act as extreme bounds representing the two phases as layers in series ($n=-1$) or parallel ($n=1$):

$$\varepsilon_{\text{comp}}^{n}=f_1\varepsilon_1^n+f_2\varepsilon_2^n+f_3\varepsilon_3^n+\cdots \tag{30}$$

As an extension of this, Lichtencker proposed the empirical logarithmic mixing rule:

$$\log\varepsilon=\sum f_i\log\varepsilon_i \tag{31}$$

The Rayleigh relation applies to a dilute dispersion of spheres:

$$\varepsilon=\varepsilon_{\text{M}}\frac{2\varepsilon_{\text{R}}+\varepsilon_{\text{M}}+2f_{\text{R}}(\varepsilon_{\text{R}}-\varepsilon_{\text{M}})}{2\varepsilon_{\text{R}}+\varepsilon_{\text{M}}-f_{\text{R}}(\varepsilon_{\text{R}}-\varepsilon_{\text{M}})} \tag{32}$$

and is equivalent to the Eshelby model and the HS lower bound (if $\varepsilon_{\text{R}}-\varepsilon_{\text{M}}$) or the upper bound (if $\varepsilon_{\text{R}}-\varepsilon_{\text{M}}$) for isotropic composites. Once again the opposite bound is found simply by interchanging the matrix and reinforcement properties. Maxwell's equation for resistivity is equivalent to Rayleigh's formulation with the dielectric constant replaced by resistivity. Self-consistent models have also been developed (Sect. 6) for which each phase is surrounded by the composite (ε) rather than phase one (ε_1). For example, for spheres (Hale 1976):

$$\varepsilon=\varepsilon_2+\frac{3f_1\varepsilon(\varepsilon_1-\varepsilon_2)}{2\varepsilon+\varepsilon_1}$$

or equivalently

$$\varepsilon=f_1\frac{\varepsilon_1-\varepsilon}{\varepsilon_1+2\varepsilon}+f_2\frac{\varepsilon_2-\varepsilon}{\varepsilon_2+2\varepsilon} \tag{33}$$

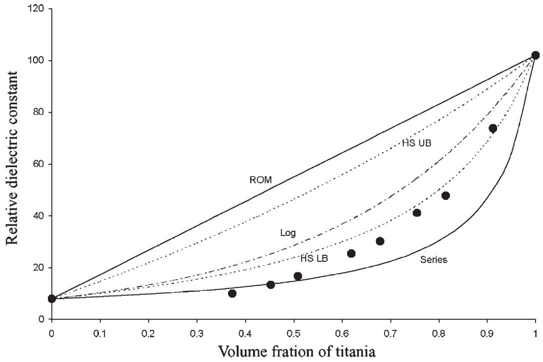


Figure 8
Plot of dielectric constant for TiO_2-fired clay composites (Kingery *et al.* 1976) alongside ROM, HS upper and lower bound, logarithmic, and series models. The Rayleigh relation and the Eshelby model are coincident with the lower bound predictions.

which is clearly symmetric with respect to both phases. All of these equations have direct analogues for electrical conductivity. Plots of dielectric constant for clay composites for the various models are shown in Fig. 8.

9. Composite Design

In attempting to identify attractive matrix/reinforcement combinations for a given application, it is often illuminating to derive "merit indices" for the performance characteristics required. These merit indices combine appropriate properties such that a single index can be maximized or minimized to optimize performance (Ashby 1993).

When considering new composite combinations, models have an important role to play. Generally, one would like to add a second phase that lies essentially perpendicular to the merit guide lines (shown as dashed lines) relative to the position of the matrix because this leads to the biggest improvements. Of course one needs to be careful to ensure that the properties of the composite do indeed lie between the extremes of the constituent phases. With regard to stiffness, for example, the composite will only have a stiffness intermediate between matrix and reinforcement if the two phases are well bonded to one another.

Predictive models give more precise information regarding the best properties that are likely to be achieved. For example, one could determine the optimal reinforcement loading/reinforcement geometry needed to obtain the target properties. Even in cases where it is not possible to make precise predictions, it is often possible to produce bounds on the best and worst properties that could arise. The framework for exploiting predictive composite models in composite

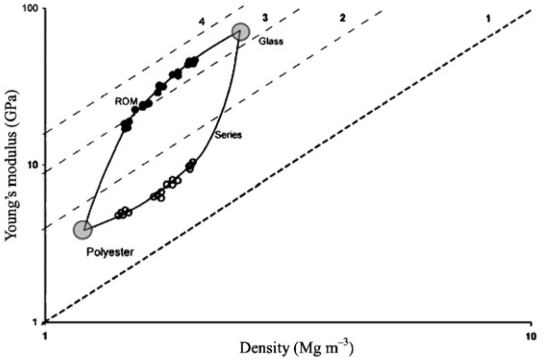


Figure 9
An elastic modulus/density selection map including bounds for the elastic modulus for polyester/glass-fiber composites. Of course on a log–log plot the ROM relation does not appear as a straight line. The dashed lines are $\sqrt{E/\rho}$ merit guidelines for designing light, but flexurally stiff, beams. Data measured parallel (filled circles) and perpendicular (open circles) to the fiber direction at various fiber fractions are also shown (Bintrup 1975) (after Ashby 1993).

design has been laid out by Ashby (1993). An example is given in Fig. 9. The stiffness/density data for glass fiber/polyester composites is well described by the upper (ROM) and lower (series) models for the axial and transverse properties respectively. Note that while better estimates than the ROM and series bounds could be employed, any difference would be relatively small on a log–log plot. In this example the axial composite performance is comparable to that of glass alone (merit index of $3\frac{1}{2}$) even at relatively low glass fractions ($<50\%$), but with the significant toughness and processing advantages characteristic of polymer composites.

Acknowledgements

I would like to acknowledge helpful discussions in this area with Dr. T. W. Clyne, Prof. L. M. Brown, Dr. O. B. Pedersen, Dr. T. Mori, and the late Dr. W. M. Stobbs, and the financial assistance of The Royal Society and the EPSRC.

Bibliography

Arsenault R J, Taya M 1987 *Acta Metall.* **35**, 651–9
Ashby M F 1993 *Acta. Metall.* **41** (5), 1313–35
Bintrup C 1975 Technische Hochschule. Aachen
Brechet Y, Embury J D, Tao S, Luo L 1991 *Acta Metall. Mater.* **39**, 1781–6
Brown L M, Stobbs W M 1971 *Phil. Mag.* **23**, 1185–99
Clyne T W, Withers P J 1993 *An Introduction to Metal Matrix Composites.* Cambridge University Press, Cambridge, UK
Cox H L 1952 *Brit J. Appl. Phys.* **3**, 72–9
Davidge R W, Green T J 1968 *J. Mater. Sci.* **3**, 629–34
Hale D K 1976 *J. Mater. Sci.* **11**, 2105–41
Hashin Z 1979 *J. Appl. Mech.* **46**, 543–50
Hashin Z, Shtrikman S 1963 *J. Mech. Phys. Solids* **11**, 127–40
Hill R 1964 *J. Mech. Phys. Solids.* **12**, 199
Hull D, Clyne T W 1996 *An Introduction to Composite Materials.* Cambridge University Press, Cambridge, UK
Hunter O, Bronwell W E 1967 *J. Am. Ceram. Soc.* **50**, 19–22
Kingery W D, Bowen H K, Uhlmann D R 1976 *Introduction to Ceramics.* Wiley, New York
Larsen D C 1979 *Property Screening of Turbine Vane Materials*, US Air Force Report, AFML-TR-79-4188
Levin V M 1967 *Mech. Solids.* **2**, 58–61
MacKenzie J K 1950 *Proc. Phys. Soc.* **B63**, 2
Mori T, Tanaka K 1973 *Acta Metall.* **23**, 571–4
Mussler B H, Shafer M W 1984 *Ceram. Bull.* **63**, 705–10
Newnham R E 1994 Nonmechanical properties of composites. In: Kelly A (ed.) *Concise Encyclopedia of Composite Materials.* Pergamon, Oxford, UK, pp. 214–20
Pedersen O B 1983 *Acta Metall.* **31**, 1795–808
Pedersen O B, Withers P J 1992 *Phil. Mag.* **65**, 1217–33
Rosen B W, Hashin Z 1970 *Int. J. Eng. Sci.* **8**, 157–73
Schapery R A 1968 *J. Comp. Mater.* **2** (3), 380–404
Selsing J 1961 *J. Am. Ceram. Soc.* **8**, 419
Shaw R R, Uhlmann D R 1987 *J. Non-Cryst. Solids.* **5**, 237
Taya M 1990 In: Andersen S I, Lilholt H, Pedersen O B (eds.) *9th Risø Int. Symp.* Risø National Laboratory, Roskilde, Denmark, pp. 201–31
Tummala R R, Firedber A L 1970 *J. Am. Ceram. Soc.* **53**, 376–80
Turner P S 1946 *J. Res. NBS.* **37**, 239
Wang X L, Hubbard C R, Alexander K B, Becher P F 1984 *J. Am. Ceram. Soc.* **77**, 1569
Weidmann G, Lewis P, Reid N 1990 *Structural Materials.* Open University, Milton Keynes, UK
Withers P J 1989 *Phil. Mag. A* **59** (4), 759–81

P. J. Withers
Manchester Materials Science Centre, UK

Composites: Interfaces

The interface between matrix and reinforcement is important in all types of composite material. The nature of the interface can influence various aspects of the performance of the composite. For example, it may be essential to ensure that load is transferred efficiently between the two constituents, so as to capitalize on the high load-bearing capacity of the reinforcement. This can be particularly important if the aspect ratio of the reinforcement is relatively low. In other cases, it may be desirable for the mechanical characteristics of the interface to be such that debonding can occur, so as to deflect propagating cracks within the material and thus to promote energy absorption and raise the toughness. A further requirement concerning the interfacial region may be that it should be highly resistant to chemical attack or to penetration of species from the environment. It is thus rather difficult to offer reliable generalizations concerning the optimal structure of interfaces in composites. Nevertheless, it is possible to identify certain principles, which can be applied when designing specific types of composite material. As a prelude to this, this article incorporates a brief outline of the technical procedures commonly used to study the mechanical characteristics of interfaces in different types of composite.

1. Mechanical Characterization of Interfaces

1.1 Interfacial Mechanics

There are two main approaches to the identification of parameters characterizing the ''strength'' of an interface in a composite. The simpler one is to assume that there is some critical stress level that will cause the interface to debond. This could be a normal stress acting transverse to the interface (crack opening, or mode I, loading) or a shear stress acting parallel to the interface (shear, or mode II, loading). In practice, both types of stress may be acting simultaneously (mixed-mode loading). It should be recognized that, whereas a crack in a monolithic material will always tend to follow a mode I path, an interface often represents a plane of weakness along which a crack will propagate even if the stress state at the crack tip is heavily mixed mode or pure shear.

One problem with critical stress level tests is that they tend to be initiation dominated. The measured stress level is usually that at which a crack is observed to propagate in an unstable manner. Such tests are therefore sensitive to the presence of flaws in the interface, at which local stress concentrations become sufficient to trigger crack growth. This initiation dependence tends to be rather marked for many composites, in which the interfaces are often relatively brittle regions. The flaws at which crack initiation occurs may be statistically representative of the structure of the interface, but they can also be created during specimen preparation. Furthermore, there is a danger with some types of test that local stress concentrations can be generated by small misalignments in the loading arrangements. It is certainly common to observe a wide scatter in measured critical stress levels for interfacial debonding. In general, a more reliable measure of interfacial ''strength'' can be obtained by adopting a fracture mechanics approach. This is based on examining the energetics of crack propagation. A crack will propagate under steady-state conditions when the driving force (expressed as an energy release rate per unit area of crack face) equals the fracture energy of the interface, which is a measure of its toughness (or strength). The initiation problem can be eliminated by ensuring that a pre-crack of suitable size is present before testing starts.

It is often a little complex to apply fracture mechanics-based tests to interfaces in composites—at least in conventional fiber composites. This is mainly because the geometry associated with a cylindrical surface means that calculation of the driving force for debonding (difference in stored elastic strain energy before and after crack advance) is not simple. Nevertheless, the disadvantages of critical stress measurements should be recognized and it may be noted that fracture mechanics analyses of several interfacial tests applicable to fiber composites have been developed. It should also be appreciated that both types of test are susceptible to effects arising from the presence of residual stresses. If the residual stress state is known, then it should be possible to introduce a correction for its effect on the apparent critical stress level or interfacial fracture energy, but in practice this effect is often overlooked or treated in a simplified manner.

Finally, a distinction should be drawn between initial debonding (crack propagation) and subsequent relative displacement of the crack flanks, i.e., frictional sliding. The latter can be of considerable importance in composites, particularly in terms of the energy absorbed during fiber pull-out—which is usually much greater than that associated with the initial debonding event. Moreover, frictional sliding usually takes place under quasisteady-state conditions, so that measured critical shear stresses for this process are expected to be more reliable than estimated critical stresses for debonding.

1.2 Test Procedures

Many test methods have been developed to interrogate the mechanical response of interfaces within composite materials. The main methods are listed in Table 1, together with an indication of some of their characteristics and identification of some references giving detailed information about them. Each test focuses on the behavior of single fibers or on the collective response of a group of fibers. In some cases, the test can be carried out on a conventional composite material, while in others it is necessary to manufacture a special specimen, commonly one containing a single isolated fiber. The table refers exclusively to fiber composites, in which the interfacial geometry is necessarily cylindrical. As a consequence of this, the interfacial stress state during testing is often predominantly mode II (shear parallel to the fiber axis).

In some of the tests, tensile stress normal to the interface is also imposed, so that mixed-mode loading is generated. (The residual stress state is often such that there is a substantial normal stress across the interface, but, since the thermal expansivities of most matrices are greater than those of most fibers, this is usually compressive and hence must be offset before a

Table 1
Classification of the tests developed for measurement of interfacial strength in fiber-reinforced composites.

Name	Entity tested	Type of specimen	Parameter measured	Loading mode	References
Fiber pull-out	Single fiber	Single embedded fiber	Critical stress or fracture energy	II	DiFrancia *et al.* (1996), Quek (1998), Zhang *et al.* (1999), Liu and Kagawa (2000)
Fiber push-out	Single fiber	Composite	Critical stress or fracture energy	II	Tandon and Pagano (1998), Yue *et al.* (1998), Kalton *et al.* (1998)
Fiber push-down	Single fiber	Composite	Critical stress or fracture energy	II	Kalinka *et al.* (1997), Kharrat and Chateauminois (1997)
Full fragmentation	Single fiber	Single embedded fiber	Critical fiber aspect ratio	II	Tripathi and Jones (1998), Shia *et al.* (2000), Park *et al.* (2000)
Protrusion (slice compression)	Group of fibers	Composite	Critical stress	II	Hsueh *et al.* (1996), Kagawa and Hsueh (1999)
Microbond	Single fiber	Single embedded fiber	Critical stress	II	Day and Young (1993), Schuller *et al.* (1998), Kessler *et al.* (1999)
Fiber bundle pull-out	Group of fibers	Composite	Critical stress or fracture energy	II	Domnanovich *et al.* (1996), Sakai *et al.* (2000)
Tensioned push-out	Single fiber	Composite	Critical stress	I/II	Watson and Clyne (1993), Kalton *et al.* (1994)
Cruciform transverse tension	Single fiber	Single embedded fiber	Critical stress	I/II	Majumdar *et al.* (1998)

mode I component to the loading can be generated.) It is thus difficult to generate pure mode I loading at a fiber–matrix interface (and, even if this were to be done, reliable detection of debonding events would be problematic in the absence of a shear stress). However, for certain types of composite, such as layered systems or coatings, the interface may be planar. Various tests have been developed for such interfaces, several of which allow mixed-mode or pure mode I loading.

2. *Interfacial Microstructures and their Effects on Mechanical Properties*

2.1 Polymer Matrix Composites

In most long-fiber polymer matrix composites (PMCs), random fiber fractures accumulate as the applied stress level is raised during axial loading. The extent of interfacial debonding, or local yielding in the interphase, determines the degree to which stress concentrations around fractured fibers become redistributed, and hence whether a catastrophic crack is likely to propagate from such fracture sites. It is therefore desirable for some kind of plastic strain accommodation to occur readily at the interface and in its immediate vicinity. However, in general interfaces in PMCs are required to be relatively strong and also to be chemically stable. For example, it is known that interfacial strength often dictates the transverse strength of a unidirectional laminate and first ply cracking of an angle-ply laminate. Furthermore, the interface must be resistant to the penetration of water and other solvents.

However, the way in which desirable properties are achieved varies somewhat, depending on the type of fiber concerned. Early development of carbon fibers relied on optimization of surface oxidation (i.e., degree of adhesion) for a good balance of properties, whereas the properties of glass fiber composites have been optimized by choice of size composition. The use of sizing resins on carbon and aramid fibers is a better way of creating an interphase, which can control the stress concentration experienced by an adjacent fiber. It should be noted that very strong interfacial bonding (relative to the strength of the fiber) is often undesirable, since, not only will it tend to make the region around fractured fibers more prone to crack propagation, but it will also tend to eliminate fiber pull-out, which is an important source of toughening in PMCs. This rather complex set of requirements for the interface in PMCs has led to the evolution of manufacturing procedures that are specific to individual composite systems. As an example of the significance of these surface treatments, it can be seen from the data in Fig. 1 that the sizing can have a substantial effect on the strain to failure under transverse loading. This property is often of critical importance in determining the resistance of the composite to crack formation during applied loading and thermal cycling. Some details of the treatments commonly applied to different fibers are given below.

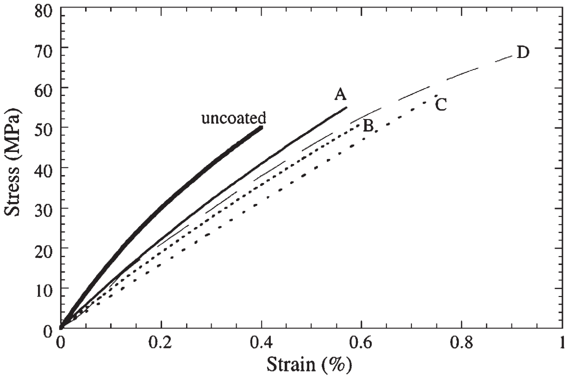


Figure 1
Stress–strain plots (Benzarti *et al.* 2001) for aligned, long fiber glass-reinforced epoxy composites, during loading in transverse tension, illustrating the effect of various fiber sizings. The sizings are: A, a commercial sizing containing bifunctional DGEBA prepolymers; B, methyl-phenyl-silicone (an elastomer, miscible but nonreacting with epoxy resins); C, a trifunctional epoxy; D, penta-glycidyl ether of cyclosilicone (reactive with the matrix and with a flexible backbone).

(*a*) *Carbon fibers.* Newly carbonized fibers do not adhere adequately to polymeric matrices. An essential stage of fiber manufacture is a surface treatment to optimize fiber–matrix adhesion. While inferior interfacial adhesion can lead to poor shear properties, over-treatment can render the composite brittle.

Carbon fibers, from acrylic or PAN precursors, are defined by the heat treatment temperature used for carbonization. High-modulus (HM) fibers tend to have a skin–core structure with a highly organized graphitic sheaf-like surface. The electrolytic oxidation employed commercially (with NH_4HCO_3 or $(NH_4)_2SO_4$ aqueous electrolytes) etches away the ends of the basal planes (which are parallel to the fiber axis) at the crystallite boundaries and functionalizes them, leading to a microporous structure exhibiting carboxylic acid, hydroxyl, amine groups, and active sites. However, with high-strength (HS) fibers, the less organized graphitic structure means that the functionalized edge sites will meet the surface at an angle.

After surface oxidation, the fibers are coated with a polymeric size, usually an epoxy resin. Whether this resin coating can be subsequently dissolved efficiently into the matrix is unclear. This aspect of the fiber–matrix interaction in carbon fiber composites is under-researched, but the possibility of a matrix–size interface also needs to be considered.

(*b*) *Glass and polymeric fibers.* For glass fiber-reinforced plastics (GRPs), the surface treatment for the fibers is more complex. Interfacial debonding is generally not observed and "matrix" yielding is the most probable interfacial event. (For high-performance polymer fibers, "interfacial" failure is likely to occur within the fiber, because of their poor transverse properties.) Surface treatments have therefore been developed that generate cross-linked polymeric coatings. It is perhaps more appropriate to describe the interfacial region as an "interphase," which may include material that has penetrated the fiber surface and has different properties from those of the matrix.

The surfaces of commercially available glass fibers (principally E-glass) have in many cases been developed specifically to facilitate composite manufacturing, with ease of handling often being at least as important as optimization of interfacial properties. The finish is applied to glass fibers immediately after water cooling at the bushing during fiber production. Typically, an aqueous emulsion consisting of the following is deposited at a rubber roller:

(i) adhesion promoter or coupling agent;
(ii) protective polymeric size or film former; and
(iii) lubricant.

Items (ii) and (iii) impart good ease of handling, and controlled wet-out kinetics with matrix resins, and are therefore chosen for compatibility during the fabrication process. For chopped fibers (chopped strand mat), an additional polymeric binder (emulsion or powder) is often deposited. For specialist requirements, such as high environmental resistance, the chemical nature of the size and binder are crucial. Item (i) is generally an organosilane, trifunctional in alkoxy groups and monofunctional in a matrix-reactive or matrix-compatible group. A common one is γ-amino propyl triethoxysilane (γ-APS). It acts to displace adsorbed water, to create a surface that can be fully wetted by the resin, and to develop strong interfacial bonds. On hydrolysis, the silane will be deposited as a complex mixture of oligomer and polymer, of which part is chemically bonded to the surface. Thus, after composite manufacture, a bonded semi-interpenetrating network of silane and matrix containing the polymeric size will form. With the diffusion of the oligomeric silane component into the matrix, a complex interphase will be created.

2.2 Metal Matrix Composites

In metal matrix composites (MMCs), the objective of introducing the reinforcement may be enhancement of stiffness, creep resistance, wear resistance, etc. It is often a prime requirement that these enhancements should be achieved without unacceptable degradation to the toughness. The scope for generating energy absorption in MMCs by processes such as fiber pull-out is usually small compared with that which is potentially available via matrix plasticity, provided the reinforcement disposition and interfacial characteristics are not such that it becomes severely inhibited or constrained. Poor bonding commonly leads to premature failure, for example by interfacial cracking or cavitation, whereas a high bond strength allows at least relatively high levels of matrix plasticity to be induced.

Another factor of significance for many types of MMC concerns interfacial chemical reaction. Processing to produce MMC materials and components commonly involves high temperatures. During such processing, the high reactivity of the metal often leads to the production of thick interfacial layers (e.g., of intermetallic compounds), which tend to be defective and brittle. These usually embrittle and weaken the composite as a whole. An example of this effect is shown in Fig. 2, which demonstrates how the tensile strength of monofilament-reinforced titanium MMCs is degraded as the interfacial reaction layer thickness increases. It is often possible to control such layer formation by the use of protective coatings on the fibers, although this is more difficult for particulate MMCs.

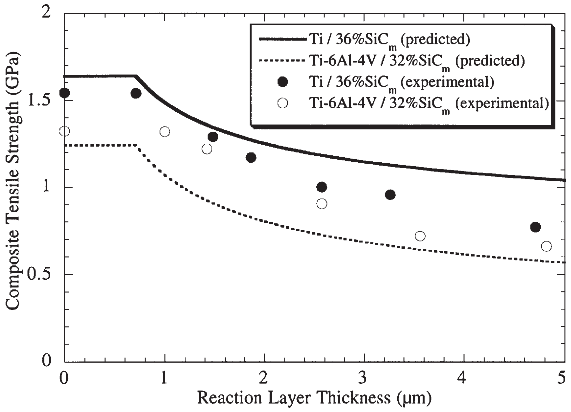


Figure 2
Experimental and predicted variation of composite strength with interfacial reaction layer thickness for axial tensile loading of two long-fiber composites—CP Ti and Ti–6Al–4V matrices reinforced with the SCS-6 monofilament (Onzawa *et al.* 1991). The predictions are obtained using a simple fracture mechanics-based model.

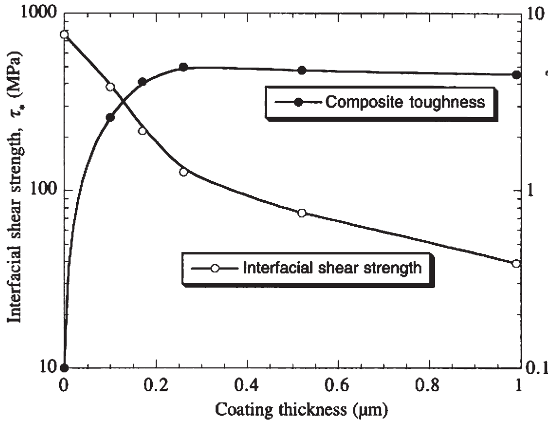


Figure 3
Effect on mechanical properties of carbon coatings on Nicalon fibers in a SiC matrix. The plot is of measured interfacial shear strength (from push-down testing) and material toughness (from the area under the stress–strain curve during loading along the fiber axis) as a function of coating thickness (Lowden 1993).

2.3 Ceramic Matrix Composites

The objective of making a composite material based on a ceramic matrix is usually to raise the toughness. This is commonly achieved by ensuring that there are weak interfaces present, at which cracks are deflected. Fabrication of such composites has always presented something of a challenge, although it is certainly easier with layered structures than with fiber reinforcement. In both cases, there is usually a requirement for extended periods at elevated temperature and this tends to cause interfaces to sinter and strengthen. It is, however, sometimes possible to introduce an interfacial coating of some sort which inhibits this. For example, graphite layers are often introduced for this purpose. Data are shown in Fig. 3 that illustrate the effects involved. The plot shows interfacial and macroscopic mechanical properties for composites made up of graphite-coated Nicalon fibers in a SiC matrix. Raising the thickness of the graphite layers leads to a decreased interfacial shear strength, but an increase in the toughness of the composite.

3. Summary

The nature of the matrix–reinforcement interface within a composite often has a profound effect on its overall properties. In general, interfaces in PMCs must be relatively strong, to promote good load transfer and to ensure that transverse properties are adequate, but not so strong as to create stress concentrations and eliminate fiber pull-out during fracture. In addition, they are often required to be chemically stable, to confer good environmental stability. It is in some cases beneficial to generate an extended interfacial region having a structure and properties different from those of the matrix—the so-called "interphase." Various chemical fiber surface treatments have been developed to achieve desirable interfacial structures in PMCs. For MMCs, the main objective is normally to ensure a strong bond, especially for particulate and short-fiber composites. Optimization of composite toughness comes from retention of as much matrix plasticity as possible. Avoidance of the formation of brittle interfacial reaction zones during processing is often a priority. Finally, in ceramic matrix composites it is usually desirable to retain weak interfaces, thus maximizing crack deflection and the associated energy absorption during mechanical loading. It is often necessary to take special measures to ensure that the interfaces remain poorly bonded during composite production.

Bibliography

Benzarti K, Cangami L, Dal Maso F 2001 Transverse properties of unidirectional glass/epoxy composites: influence of fibre surface treatments. *Composites A* **32**, 197–206

Clyne T W 1996 The effect of interfacial characteristics on the mechanical performance of particulate, fibrous and layered metal matrix composites—a review of some recent work. *Key Eng. Mater.* **127–31**, 81–98

Day R J, Young R J 1993 Raman microscopy of fibers and composites. *J. Microscopy* **169**, 155–61

DiFrancia C, Ward T C, Claus R O 1996 The single-fibre pull-out test: 1. Review and interpretation. *Composites A* **27**, 597–612

Domnanovich A, Peterli H, Kromp K 1996 Determination of interface parameters for carbon/carbon composites by the fibre-bundle pull-out test. *Compos. Sci. Technol.* **56**, 1017–29

Dorey G 1990 Fibre–matrix interface and composite strength. In: Middleton D H (ed.) *Composite Materials in Aircraft Structures*. Longman, pp. 50–66

Drzal L T, Herrera-Franco P J, Ho H 2000 Fibre–matrix interface tests. In: Carlsson L, Crane R L, Uchino K (eds.) *Comprehensive Composite Materials: Vol. 5. Test Methods, Nondestructive Evaluation and Smart Materials*. Elsevier, Amsterdam, pp. 71–111

Hseuh C H 1994 Slice compression tests versus fiber push-in tests. *J. Compos. Mater.* **28**, 638–55

Hsueh C H, Brandon D G, Shafry N 1996 Experimental and theoretical aspects of slice compression tests. *Mater. Sci. Eng. A* **205**, 91–100

Hull D, Clyne T W 1996 *An Introduction to Composite Materials*. Cambridge University Press, Cambridge pp. 133–57

Jones F R (ed.) 1996 *Handbook of Polymer–Fibre Composites*. Longman, London, pp. 29–48, 231–44, 265–85

Jones F R 1996 Interphase formation and control in fibre composite materials. *Key Eng. Mater.* **116**, 41–59

Kagawa Y, Hsueh C H 1999 Model experiment on a protrusion method for measurement of interface shear sliding stress in fiber-reinforced composite. *Mater. Sci. Eng. A* **271**, 70–8

Kalinka G, Leistner A, Hampe A 1997 Characterisation of the fibre/matrix interface in reinforced polymers by the push-in technique. *Compos. Sci. Technol.* **57**, 845–51

Kalton A F, Howard S J, Janczak-Rusch J, Clyne T W 1998 Measurement of interfacial fracture energy by single fibre push-out testing and its application to the titanium–silicon carbide system. *Acta Mater.* **46**, 3175–89

Kalton A F, Ward-Close C M, Clyne T W 1994 Development of the tensioned push-out test for study of fibre–matrix interfaces. *Composites* **25**, 637–44

Kessler H, Schuller T, Beckert W, Lauke B 1999 A fracture-mechanics model of the microbond test with interface friction. *Compos. Sci. Technol.* **59**, 2231–42

Kharrat M, Chateauminois A 1997 On the interfacial behaviour of a glass/epoxy composite during a micro-indentation test: assessment of interfacial shear strength using reduced indentation curves. *Composites A* **28**, 39–46

Kim J K, Mai Y W 2000 Effect of interface strength on MMC properties. In: Clyne T W (ed.) *Comprehensive Composite Materials: Vol. 3. Metal Matrix Composites*. Elsevier, Amsterdam, pp. 117–49

Lesko J J, Jayaraman K, Reifsnider K L 1996 Gradient interphase regions in composite systems. *Key Eng. Mater.* **116**, 61–86

Lewis M H 2000 Interfaces in ceramic matrix composites. In: Warren R (ed.) *Comprehensive Composite Materials: Vol. 4. Ceramic Matrix Composites*. Elsevier, Amsterdam, pp. 289–322

Liu H Y, Mai Y W 1999 An appraisal of composite interface mechanics models and some challenging problems. *Compos. Interfaces* **6**, 343–62

Liu Y F, Kagawa Y 2000 The energy release rate for an interfacial debond crack in a fiber pull-out model. *Compos. Sci. Technol.* **60**, 167–71

Lowden R A 1993 Fibre coatings and the mechanical properties of a continuous fiber reinforced SiC matrix composite. In: Peteves S D (ed.) *Designing Ceramic Interfaces II*. Commission of European Communities, Luxembourg, pp. 157–72

Mackin T J, Warren P D, Evans A G 1992 Effects of fibre roughness on interface sliding in composites. *Acta Metall. Mater.* **40**, 1251–7

Mader E 1997 Study of fibre surface treatments for control of interphase properties in composites. *Compos. Sci. Technol.* **57**, 1077–88

Majumdar B S, Gundel D B, Dutton R E, Warrier S G, Pagano N J 1998 Evaluation of the tensile interface strength in brittle-matrix composite systems. *J. Am. Ceram. Soc.* **81**, 1600–10

Naslain R R 1998 The design of the fibre–matrix interfacial zone in ceramic matrix composites. *Composites A* **29**, 1145–55

Onzawa T, Suzumura A, Kim J H 1991 Influence of reaction zone thickness on tensile strength for titanium matrix composites reinforced with SiC fiber. In: Tsai S W, Springer G S (eds.) *Composites: Design, Manufacture and Application.* SAMPE, HI, pp. 19J/1–19J/10
Park J M, Kim J W, Goda K 2000 A new method of evaluating the interfacial properties of composites by means of the gradual multi-fiber fragmentation test. *Compos. Sci. Technol.* **60**, 439–50
Piggott M R 1997 Why interface testing by single-fibre methods can be misleading. *Compos. Sci. Technol.* **57**, 965–74
Quek M Y 1998 The single fiber pull-out test analysis: influence of a compliant coating on the stresses at bonded interfaces. *J. Adhesion Sci. Technol.* **12**, 1391–412
Sakai M, Matsuyama R, Miyajima T 2000 The pull-out and failure of a fiber bundle in a carbon fiber reinforced carbon matrix composite. *Carbon* **38**, 2123–31
Schuller T, Bahr U, Beckert W, Lauke B 1998 Fracture mechanics analysis of the microbond test. *Composites A*. **29**, 1083–9
Shia D, Hui C Y, Phoenix S L 2000 Statistics of fragmentation in a single-fiber composite under matrix yielding and debonding with application to the strength of multi-fiber composites. *Compos. Sci. Technol.* **60**, 2107–28
Tandon G P, Pagano N J 1998 Micromechanical analysis of the fiber push-out and re-push test. *Compos. Sci. Technol.* **58**, 1709–25
Tripathi D, Jones F R 1998 Single fibre fragmentation test for assessing adhesion in fibre reinforced composites. *J. Mater. Sci.* **33**, 1–16
Watson M C, Clyne T W 1993 The tensioned push-out test for measurement of fibre/matrix interfacial toughness under mixed mode loading. *Mater. Sci. Eng* **A160**, 1–5
Yue C Y, Lee L L, Sano T 1998 Experimental push out testing and analysis of fibre reinforced composites: applicability and test considerations. *Mater. Sci. Technol* **14**, 987–1000
Zhang X, Liu H Y, Mai Y W, Diao X X 1999 On steady-state fibre pull-out: I. The stress field. *Compos. Sci. Technol* **59**, 2179–89

T. W. Clyne[a] and F. R. Jones[b]
[a] *University of Cambridge, UK*
[b] *University of Sheffield, UK*

Concrete as a Building Material

Concrete is the most widely used engineering material in the world, with an annual consumption of around 5 billion tonnes. Concrete structures, including buildings, bridges, power plants, dams and other hydraulic structures, constitute a large part of the modern civil infrastructure. The major constituents of concrete are Portland cement, aggregates, and water. When mixed together, the cement reacts with water to form a binder that holds the aggregates together. The resulting composite material is strong in compression but relatively weak in tension. Therefore, steel reinforcements are often incorporated to carry the tensile stresses. Since reinforced concrete is covered elsewhere in this encyclopedia, the present article will focus on plain concrete alone (see *Reinforced Concrete**).

The typical properties of concrete are listed in Table 1, together with those for steel, its major competitor for the construction of large structures. It is obvious that steel is far superior in terms of strength and modulus. However, the low cost of concrete makes it the material of choice in many cases. Moreover, the energy content of concrete (i.e., the energy consumed in producing a unit volume of the material) is much lower. Its use over steel is therefore favorable from an energy and environmental conservation viewpoint. As a construction material, concrete possesses a number of other advantages. Since concrete stays plastic for some time after the constituents are mixed, members of complex shapes can be easily constructed by casting. The shape flexibility makes possible the minimization of joints, and can enhance the aesthetics of structures. With high resistance to both water and fire, concrete is widely used in dams and hydraulic structures as well as in buildings. In building applications, concrete also offers good thermal insulation that serves to reduce the energy cost for heating/cooling.

1. *Constituents of Concrete*

Cement is the binding phase in concrete and it usually constitutes about 15 to 25% of the concrete weight. Details on cement production, composition, hydration, and microstructures are provided in a separate article in this encyclopedia (see *Portland Cement**). Here, it suffices to mention that the cement content in the concrete, and the water/cement ratio, are important parameters governing the properties of concrete at the fresh and hardened states, as well as durability in the long term. Aggregates in concrete are classified into coarse and fine aggregates, with the boundary set arbitrarily at about 4.75 mm (3/16 in), which is the size of a No. 4 sieve. For fine aggregates, the smallest size is often around 0.075 mm. For reinforced concrete structures, the maximum aggregate size seldom goes beyond 25 to 40 mm, as the aggregates need to be smaller than the reinforcement spacing by a certain margin. For mass concrete, aggregates up to 150 mm in size are commonly used.

In most concrete constructions, aggregates are obtained from local sources. Indeed, any natural or crushed aggregate can be used in concrete provided the following requirements are satisfied. The aggregate should be strong, dense, and difficult to cleave. Also, it should not react with alkalis (e.g. NaOH or KOH) in cement. Aggregates contaminated with deleterious materials, such as organic impurities or salt, need to be washed until they are clean. The soundness of aggregates is important as well. In other words, the aggregate should be dimensionally stable under

Table 1
Comparison between concrete and steel as construction materials.

Material	Concrete	Steel
Density ($\mathrm{kg\,m^{-3}}$)	2400	7800
Strength (MPa)	3–6 (tension) 30–100 (compression)	250 to >1000
Failure strain (%)	~ 0.01 (tension) 0.1–0.2 (compression)	10 to 35
Young's modulus (GPa)	30–50	210
Energy requirement ($\mathrm{GJ\,m^{-3}}$)	3	300
Energy content ($\mathrm{kWh\,m^{-3}}$)	400–800	~ 20000
Cost (\$ $\mathrm{tonne^{-1}}$)	20–50	250 to >1000

environmental changes such as freezing/thawing and wetting/drying.

Most natural aggregates have a density between $1520\,\mathrm{kg\,m^{-3}}$ and $1680\,\mathrm{kg\,m^{-3}}$, and are referred to as normal weight aggregates. The density of concrete produced with these aggregates is around $2400\,\mathrm{kg\,m^{-3}}$. The strength of normal weight aggregates ranges from 70 to 350 MPa. This is sufficient for common applications. However, when high-strength concrete (with strength up to or above 100 MPa) is to be produced, the engineer needs to ensure that strong aggregates are employed. In some cases, to reduce the loading on the foundation or supporting structures, it is desirable to reduce the dead weight of the superstructure. Lightweight aggregates, with a density of below $1120\,\mathrm{kg\,m^{-3}}$, are used to produce concrete with a density below $1850\,\mathrm{kg\,m^{-3}}$. Lightweight aggregates can be obtained from natural sources (e.g., volcanic tuff or pumice) or produced artificially through thermal treatments of clays, slag, fly ash, etc. Heavyweight aggregates are defined as those with a density above $2080\,\mathrm{kg\,m^{-3}}$. Examples include barium ore, iron ore, or steel shots. With the capability to attenuate gamma rays, x-rays, and neutrons effectively, heavyweight aggregates are used in radiation shields for hospitals and nuclear power plants.

Water for making concrete should ideally be suitable for drinking. Impurities in water may affect the setting of cement, reduce the strength of concrete, cause staining of the concrete surface, or lead to corrosion of the steel reinforcement. In reality, the use of undrinkable water is not necessarily harmful. When in doubt, setting time and strength tests can be carried out on samples made with both clean water and the water of unknown quality. Note that some impurities such as salt do not affect concrete properties but can significantly accelerate steel corrosion. Seawater should therefore never be used.

Besides water, cement, and aggregates, other materials may be added to concrete for various purposes. These materials are collectively called admixtures, and may be broadly categorized into surfactants, set-controlling admixtures, and mineral admixtures. Surfactants are long-chain organic molecules adsorbed to the surface of cement particles. Some surfactants are air-entraining agents as their presence will lower the surface tension, leading to formation of small bubbles around solid particles. With a uniformly distributed system of small bubbles, the freeze/thaw resistance of concrete can be greatly improved. Other surfactants, called water-reducing admixtures, introduce charges on the surface of cement particles, causing repulsion and dispersion. As a result, water trapped inside flocculated cement particles is released, thus improving the flow properties of concrete. A surfactant that is particularly effective in water reduction (several times more effective than normal water-reducing agents) is called a superplasticizer, and it is a key component in the production of high-strength and high-workability concrete.

Set-controlling admixtures are added to retard or accelerate the hydration reaction between water and cement. When concreting is carried out in hot weather, the chemical reaction is so fast that setting may occur very rapidly, allowing little time for proper placing and finishing. Also, the rapid generation of hydration heat may lead to degradation in long-term strength. The addition of a set-retarder can alleviate the situation. Conversely, during cold-weather concreting, to attain sufficient strength within a reasonable period of time, an accelerator is required.

Surfactants and set-controlling admixtures are often added to the cement in small amounts (<1 wt.%), and their addition has no effect on the final hydration products. Mineral admixtures, on the other hand, are applied in much larger quantities (over 10 wt.%). Mineral admixtures are available in nature, with the volcanic glasses used in Roman concrete being classical examples. However, most mineral admixtures employed today, such as fly ash, blast furnace slag, and silica fume, are byproducts of industrial processes. The use of such admixtures in concrete therefore also provides the solution to a waste disposal problem. Many common mineral admixtures are pozzolans. By definition, these are materials that possess no cementitious properties on their own, but can react with calcium hydroxide (a reaction product of cement hydration) to form cementitious materials. Examples include low-calcium fly ash and silica fume. Others, such as blast-furnace slag, can react directly with water to form a cementitious material. High-calcium fly ash is both pozzolanic and cementitious. Due to their cementitious or pozzolanic properties, mineral admixtures are often used to replace part of the cement in the concrete. By reducing the cement content, heat of hydration is reduced, thus lowering the chance of thermal cracking in thick structures. When fly ash or blast furnace slag is used to replace cement, the early strength is usually reduced, but the ultimate strength is higher than that of concrete without cement replacement. For silica fume, its small size allows the pozzolanic reaction to occur very rapidly, resulting in higher early strength as well. With the use of mineral admixtures, the final microstructure of the concrete becomes denser and less permeable. Both the strength and durability of concrete are thus improved.

2. *Concrete at Fresh State*

After the constituents are mixed, concrete initially behaves as a fluid and can be cast to form the required shape. As the cement hydration reaction progresses, concrete will eventually change from a fluid to a rigid state. This stiffening process is referred to as setting. Stiffening is indeed a gradual process, but two reference points in the process, referred to as initial set and final set, have been defined. These points are obtained through a standard penetration resistance test. Physically, the initial set represents the approximate time when the concrete can no longer be properly mixed and compacted. The final set is the approximate time when concrete strength starts to develop at a high rate. In engineering applications of concrete, knowledge of initial and final set values can allow the proper scheduling of various operations, such as mixing, placing, and formwork removal.

To make concrete components of good quality, the fresh concrete must exhibit good workability. Workability can be defined as the property governing the required effort to place, compact, and finish a concrete mix with minimum loss of homogeneity. It can be described in terms of two separate components, consistency (the ease of flow) and cohesiveness (the tendency for segregation). Workability, as formally defined, is very difficult to measure. For practical applications a number of simple tests, including the slump test, Vebe test, compacting factor test, flow table test, and the ball penetration test, have been developed. The ASTM and BS publications contain details of the tests (see, for example, Klieger and Lamond 1994). It should be noted that the above tests attempt to measure consistency only. Cohesiveness is often assured through empirical knowledge regarding the proportioning of concrete mix, as well as proper procedures during concrete transportation and placing to prevent segregation.

The workability of concrete is affected by many factors. By increasing the water content, consistency is improved. For example, by keeping the same water/cement ratio but increasing the cement/aggregate ratio, the mix exhibits higher workability. For the same water/cement/aggregate ratio, consistency can be improved by using coarser sand, and by replacing rough aggregates with round/smooth ones. When the consistency of the mix becomes too high (e.g., through the excessive use of water), the components tend to separate, and the cohesiveness is adversely affected. Care should therefore be taken to achieve an optimum. In some cases, cohesiveness and consistency can be improved at the same time. Examples include the partial replacement of fine aggregates with small fly ash particles, and the addition of silica fume to very low water/cement ratio mixes. Besides material parameters, the type of construction also affects workability. For example, a mix may have sufficient workability for a lightly reinforced concrete slab, but is not workable for the heavily reinforced wall of a power plant. Workability also varies with time. At a longer time after mixing the consistency becomes lower due to the loss of water through absorption, evaporation, and chemical reaction. This reduction in consistency is commonly referred to as slump loss. It occurs most commonly in hot weather, and can be prevented by using ice to replace part of the mixing water.

In most concrete constructions the concrete is batched and mixed in a central plant, and transported to the site. During transportation, agitation is applied to prevent segregation. After arrival at the site, concrete is placed as near as possible to its final position. Compaction is achieved through the use of internal immersion-type vibrators or external form-type vibrators. During vibration, it is common for water to separate from the other components, and travel upwards. This phenomenon is known as bleeding. Most bleed water gets to the concrete surface, but some will be trapped under large aggregate particles or the reinforcing bars to form weak zones with an increased water/cement ratio. Some methods to reduce bleeding include the use of finer cement particles, addition of fine pozzolans or air entraining agents, and the reduction of water content (provided sufficient workability can be maintained).

After casting and compacting, concrete needs to be properly cured for strength development. The objective of curing is to keep concrete wet until the hydration reactions have progressed to a certain level. It is usually achieved by water spraying, ponding, or by covering concrete with wet sand, hessian, or burlap. Curing membranes, which are plastic sheets applied on the concrete surface, can also be used. In most cases, curing needs to be carried out for seven days. At lower temperatures, longer time is required. If curing is not properly carried out, the ultimate strength of concrete can be significantly affected. Moreover, if there is excessive surface water evaporation when concrete is still in the plastic state, plastic shrinkage will occur. In many cases, plastic shrinkage is accompanied by surface cracking. Besides being unsightly, the cracks often run on top of steel reinforcements, and act as convenient paths for water to reach the steel. The durability of the concrete structure will then be affected.

Besides site casting, concrete members can also be precast in the factory and transported to the site for assembly. This approach is advantageous when the same type of component is repeatedly used in different parts of a structure, or in a large number of identical structures. In precasting, concrete is often placed and compacted in the same way as in site casting. Research has demonstrated the possibility of fabricating concrete components through a continuous extrusion process. The extrusion approach is particularly suitable for nonstructural components such as tiles, pipes, roofing, and partition walls. To improve the properties of extruded products, small fibers are often added. In precasting work, the turnover rate is an important consideration, and it is desirable to accelerate the development of strength so that the final product can be delivered to the site at the earliest possible time. Steam curing is the common approach to obtain high strength at an early age. Various researchers have investigated the use of microwave energy to accelerate concrete curing, and the results appear to be promising.

3. *Hardened Concrete Under Loading*

The mechanical behavior of concrete depends on properties of the hardened cement paste, the aggregates, and the interphase between the aggregate and the paste. The paste–aggregate interphase is a thin layer of cementitious material (about 10–20 μm in thickness) around the aggregate that possesses a higher porosity than the bulk paste. Its formation is attributed to ineffective particle packing at the aggregate surface (compared with the bulk), as well as the trapping of bleed water at the bottom of the aggregates. For normal weight concrete, the interphase is weaker than both the aggregate and the bulk paste. Also, when loading is applied, stress concentrations exist around aggregates due to the mismatch in paste and aggregate modulus. Low strength of the interphase, together with the presence of stress concentration, makes it the weakest link at which damage will start.

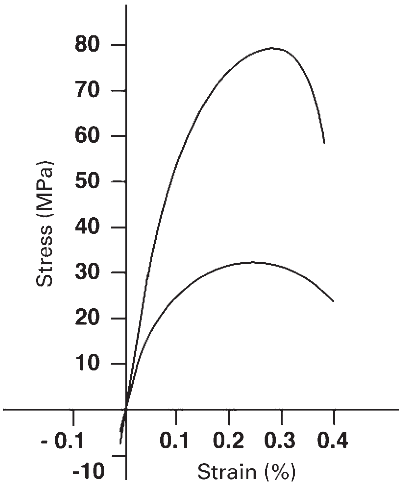


Figure 1
Stress–strain curves for normal and high-strength concrete (positive sign is used for compressive stress and strain).

The stress–strain curves for two different types of concrete are shown in Fig. 1. Compressive behavior is shown on the positive quadrant. The lower curve represents normal-strength concrete, while the upper one is an example of high-strength concrete. For both types of concrete, the compressive behavior is approximately linear up to about 30–40% of the ultimate strength (for normal- and high-strength concrete, respectively). Nonlinear behavior then commences due to the formation of microcracks within the interphase. When compressive loading is gradually increased beyond the linear range, the microcracks initially propagate in a stable manner, first along the aggregate surfaces, and then into the cement paste. On further loading, some of the microcracks coalesce into localized macrocracks. Unstable propagation of the macrocracks leads to the "crushing" of concrete (or disintegration into multiple fragments) when the compressive strength is reached.

Many features about the failure of concrete can be explained in terms of the cracking process. Since repeated loading can accelerate the propagation of cracks, concrete strength under cyclic loading is lower than the static strength. Under a given loading, cracks can grow to larger sizes if sufficient time is allowed for the propagation to occur. Therefore, if concrete is stressed to a high level (above 75% of the ultimate load) and the loading is maintained, failure will eventually occur. Since crack formation and propagation is accompanied by an increase in volume, the presence of confining stress can delay and stabilize crack growth. As a result, both the compressive strength and ductility of concrete increase with lateral confinement.

When concrete is under tension, the growth and coalescence of microcracks occur much more easily.

Crack localization and tensile failure therefore occur at a stress significantly lower than the compressive strength. An empirical formula relating the tensile strength (f_t) and compressive strength (f_c) is given by:

$$f_t = 0.615(f_c)^{0.5} \quad (\text{for } 21\ \text{MPa} < f_c < 83\ \text{MPa})$$

Substituting numerical values for f_c, f_t is found to be around 7–13% of the compressive strength, with a lower f_t/f_c ratio for higher concrete strength. In the above formula, f_c is obtained from the direct compression of cylinders while f_t is measured by the splitting tensile test. In some countries, compressive strength is obtained from the cube specimen instead. Due to the effect of lateral constraint from the top and bottom platens, the cube strength can be higher than the cylinder strength by 5–25%. Normally, the two measurements become closer as concrete strength is increased. To measure the tensile behavior of concrete, the flexural test is also commonly performed on a standard beam specimen. Depending on the concrete mix, the flexural strength is 50–100% higher than the splitting strength.

The water/cement ratio (w/c), or water/binder ratio when mineral admixtures are employed, is the most important factor governing concrete strength. The variation of f_c with w/c is shown in Fig. 2, where the two curves represent the upper and lower bounds. As cement hydration progresses, the space originally occupied by water is filled with hydration products. For a lower w/c, a relatively large amount of hydration products is available to fill a smaller space. The hardened cement paste (both in the bulk and within the interphase) becomes denser. Hence it is more difficult for crack initiation and growth to occur. While it is possible to make high-strength concrete by reducing w/c alone, it is often more effective to increase concrete strength with the addition of pozzolanic admixture as well. The pozzolanic reaction converts calcium hydroxide, the weak phase in the hardened paste, into C–S–H gel, the major cementing material. Also, when very fine particles such as silica fume is added, microvoids are filled and the compactness of the material is improved. It should be noted that concrete workability is significantly reduced when w/c becomes very low. The addition of superplasticizer is then essential to maintain the proper rheological properties.

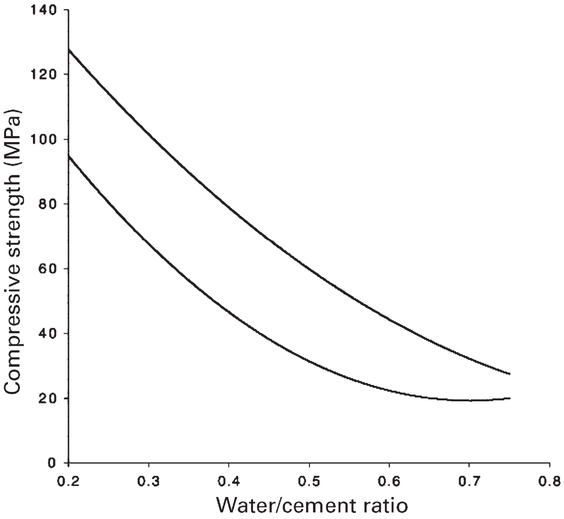


Figure 2
Variation of concrete strength with water/cement ratio (lines represent upper and lower bounds).

Other factors affect concrete strength besides w/c. If the aggregate is weaker than the hardened cement paste, concrete strength will be limited by aggregate failure. Rough and angular aggregates make stronger concrete than round and smooth ones, as it is more difficult for cracks to go around them. With larger aggregates, the size of the weak interfacial zone is also increased, resulting in reduced strength. For the production of high-strength concrete the maximum aggregate size is therefore limited to 19 mm. The curing process also has an important effect on strength. High temperature at the initial curing stage leads to more rapid strength gain, but the long-term strength is reduced. For concrete with admixtures, the amount of admixture added as well as its type (silica fume, fly ash, slag, etc) strongly affects the final strength. The effects of these other variables are reflected in the width of the "band" in Fig. 2, which shows the variation of f_c for a given w/c.

The elastic modulus of concrete is usually determined from test data up to about 50% of the ultimate strength (i.e., within or slightly beyond the linear range). From Fig. 1, it can be observed that high-strength concrete also has a higher modulus. In high-strength concrete, the w/c is lower, and hence both the hardened paste and the interphase are denser and stiffer. Also, to make high-strength concrete, good quality aggregates with lower porosity are often employed. With all the components being stiffer, the composite modulus of concrete is increased.

In the design of most reinforced concrete structures, the low tensile strength of concrete is neglected and steel reinforcement is relied upon to carry the tensile load. However, in a number of situations, such as the failure of anchoring elements in concrete, bond failure of steel reinforcements, cracking in concrete dams, failure of unreinforced concrete pipes as well as shear and torsional failure of concrete members, the structural capacity is strongly dependent on the tensile properties of the concrete. In these cases, the governing material parameters include not only the tensile strength, but also the postpeak tension-softening curve (Fig. 3). Physically, the softening behavior results from the bridging of cracks by aggregates, and it therefore reflects the ability of the

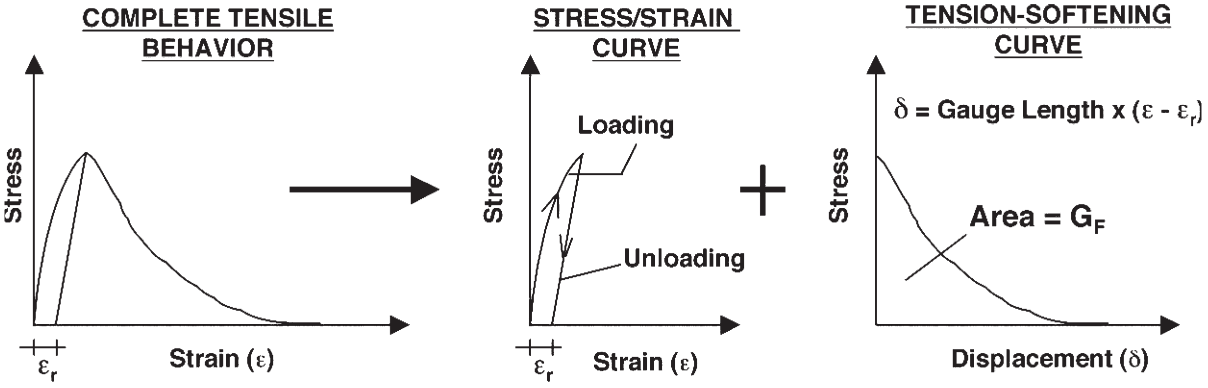


Figure 3
Derivation of the tension-softening curve from the complete tensile behavior.

material to resist crack propagation. The study of tension softening and its proper implementation in computational analysis have been two major foci in the investigation of the fracture mechanics of concrete.

Currently, computational tools are available to simulate structural behavior when failure is dominated by the tensile behavior of concrete (e.g., as in the situations mentioned above). Moreover, with improved understanding of tensile failure in concrete, it is now possible to explain and quantify the well known "size effect" in concrete structures. The capability to quantify this phenomenon of decreasing load carrying capacity with increasing member size is of great engineering significance.

The area under the tension-softening curve represents the total energy required to create a unit area of crack surface, and is called the fracture energy (G_F). A characteristic length, l_{ch}, defined as EG_F/f_t^2, where E is Young's modulus, is an useful parameter reflecting the brittleness of concrete. With a lower value of l_{ch}, failure will occur in a more brittle manner. As concrete strength is increased, f_t increases at a faster rate than E and G_f. As a result, failure of high-strength concrete occurs in a more catastrophic manner than normal strength concrete.

The incorporation of fibers is an effective way to improve the crack resistance of concrete and reduce its brittleness. Fibers made of steel, glass, polypropylene, carbon, PVC, and other polymers are commercially available. With fiber addition, besides the increase in fracture energy, the flexural strength, shear strength, fatigue resistance, and impact resistance of concrete can all be increased. More importantly, the "ductility" of concrete, indicated by the deformation capacity before final failure, is greatly improved. Indeed, when the fiber volume fraction goes beyond a critical value, concrete will show a pseudo-strain-hardening behavior under tension: after first cracking, stress will continue to increase with strain. Physically, this is due to the high bridging stress provided by the fibers that makes possible the formation of multiple fine cracks in the member as loading is increased beyond first cracking. Pseudo-strain-hardening concrete has very high shear strength and energy absorption as well as excellent crack control. Fibers are therefore very useful in improving the durability and seismic resistance of concrete structures.

When concrete is kept under sustained loading, creeping occurs and the deformation increases with time. Many theories (e.g., moisture diffusion, gel sliding, microcracking, etc.) have been proposed to explain the mechanism of concrete creep. However, no universal theory that can explain all the experimental observations is yet available. Since creeping occurs in the cement paste only, it can be reduced with the use of stiffer aggregates, or a higher aggregate fraction. Being a thermally activated process, creep increases with temperature. When air humidity is low, or when the concrete member has a large surface/volume ratio, there will be increasing moisture movement from the concrete to the surrounding environment, resulting in more creep. The creep of concrete is also affected by the density (or gel/space ratio) of the hydrated cement. With lower w/c and higher age (or higher maturity) when loading is applied, creep is reduced.

Time-dependent dimensional change of concrete can also occur under zero applied loading. When moisture moves out of concrete drying shrinkage occurs. Since both drying shrinkage and creep are dependent on moisture movement and are governed by the behavior of the cement paste, many parameters affect both in a similar manner. Shrinkage also occurs as carbon dioxide reacts with calcium hydroxide in the paste to from calcium carbonate. Since carbonation proceeds slowly from the surface of the concrete, carbonation shrinkage only occurs in the surface layer. Its effect, however, is implicit in many experimental data for concrete shrinkage. Under many common situations, concrete shrinkage is restrained and tensile stresses are induced. To prevent excessive shrinkage cracking, reinforcements in the form of rods or fibers are incorporated for crack control.

In practice, design charts are available for the prediction of creep and shrinkage. Note that shrinkage occurs simultaneously with creep when the concrete member is under loading. The effect of shrinkage is therefore implicit in the charts for creep.

4. Other Important Properties of Concrete

The thermal properties of concrete are of interest in that nonuniform or restrained thermal expansion and contraction induce stresses. The thermal conductivity and heat capacity of concrete govern the insulation and energy efficiency of buildings with concrete walls. In concrete buildings, fire resistance is an important concern. The expansion coefficient (α) of concrete depends on the aggregate type, the mix proportion, the age of the concrete, and moisture content. For both normal and lightweight concrete, α ranges from 7.5 to $13 \times 10^{-6}\,\mathrm{K}^{-1}$. The thermal conductivity of concrete increases with material density (or, decreases with porosity). The value ranges from $1.5\text{–}3.5\,\mathrm{W m^{-1} K^{-1}}$ for normal-weight concrete to $0.5\text{–}1.1\,\mathrm{W m^{-1} K^{-1}}$ for lightweight concrete, which is more porous. The heat capacity of concrete is insensitive to the aggregate type, but is dependent on the porosity (or w/c) and moisture content of the cement paste. Its value normally ranges from 800 to $1200\,\mathrm{J\,kg^{-1} K^{-1}}$. In the case of fire, the high heat capacity delays the heating-up of concrete members. Also, the low thermal conductivity delays the penetration of heat into the members. As temperature increases, the strength of concrete is reduced. However, up to 500 °C concrete can still retain up to 75% of its room temperature strength. The fire resistance of concrete is better than steel and timber, the other two major construction materials. Indeed, concrete is often employed in buildings for the fire protection of steel members.

The electrical properties of concrete are useful parameters for the nondestructive evaluation of concrete structures. The resistivity of the concrete cover is an indicator of the probability of steel corrosion, as low resistivity limits the flow of ions around the steel and hence slows down electrochemical reactions. When radar is employed to measure the thickness of concrete members, or to identify delaminations, the wave velocity depends on the dielectric constant of concrete, while the depth of penetration is a function of both the dielectric constant and conductivity. Values of these parameters are therefore required for the choice of radar wavelength and proper interpretation of results.

5. Durability of Concrete

To design durable concrete structures, a good understanding of the deterioration mechanisms of concrete is essential. Concrete may degrade under physical action, chemical action, or a combination of the two. An example of physical action is the freezing of water in saturated capillary pores. The volume increase due to ice formation induces high local stresses and microcracking. With repeated freezing and thawing, damage becomes more severe. To protect concrete from freeze-thaw damage, air-entraining agents are used to introduce small bubbles (0.05 to 1 mm in size) at close spacing (about 0.2 mm). When ice starts to form in the capillary pore, the unfrozen water, which is under pressure, moves to the bubbles and is temporarily stored. With part of the water moving out of the capillary pore, freezing no longer introduces high stresses. Air entrainment increases the overall porosity of concrete but this reduces the strength for a given w/c. However, the incorporation of air bubbles improves the consistency of concrete. Keeping the same workability, the strength reduction can be partly compensated for by using a lower w/c.

Concrete degrades in the presence of sulfates (e.g., from groundwater). Calcium sulfate reacts with monosulfate hydrate (a hydration product of calcium aluminate in the cement) to form an expansive reaction product. The physical effect of expansion results in disruption of the concrete. Sodium sulfate can also react with the calcium hydroxide; with magnesium sulfate the C–S–H gel is attacked as well. The chemical conversion of calcium hydroxide and C–S–H into calcium sulfate reduces the strength and stiffness of the paste. The severity of sulfate attack therefore depends on the cation and increases in the order calcium—sodium—magnesium. To control sulfate attack, cement with a low percentage of calcium aluminate can be used. As the chemical reactions involve water, a low w/c can be specified to achieve lower concrete permeability. The replacement of cement with pozzolans is also effective as a means to remove the hydroxide, "dilute" the aluminate, and decrease the concrete permeability.

When the aggregates of concrete contain certain minerals (such as glassy silicates or dolomite), and the cement contains a relatively high alkali content ($>0.6\%$ sodium equivalent), alkali–aggregate reaction will occur. The absorption of water, by either the reaction product or minerals exposed by the reaction, leads to expansion and internal stresses. To prevent alkali–aggregate reaction, cement with low alkali content or nonreactive aggregates should be employed. In cases where this is not practical, the concrete surface can be coated to prevent water penetration.

The corrosion of steel reinforcements in concrete structures is a major durability problem. When the pH inside concrete is high, a stable iron oxide layer is formed on top of the steel, and corrosion is not a concern. However, as carbonation (the reaction between carbon dioxide and calcium hydroxide) reaches the level of the steel, the pH drops to a value low enough for the protective layer to break down. Corrosion can then occur at a significant rate. The protective layer will also be disrupted when a critical chloride ion concentration is reached. For coastal structures and pavements de-iced with salt, penetration of chloride ions can therefore lead to steel rusting. Steel corrosion leads to the formation of reaction products that can absorb water and expand to several times the size of the reacted iron. The associated stress often leads to spalling of the concrete cover. Once steel is exposed, the corrosion rate can increase greatly. To prevent steel corrosion, sufficient concrete cover should be provided and a low w/c employed to slow down the penetration of water, gas, and ions. In cases where corrosion has already started, cathodic protection, as well as realkalization and desalination, can be carried out.

Since many degradation mechanisms involve the transport of corrosive agents into concrete, it is always desirable to reduce the water permeability and gas/ion diffusivities of concrete by the use of a low w/c. Indeed, if the w/c is reduced from 0.6 to 0.4, concrete permeability decreases by about 20 times. However, when the w/c is reduced it is important to maintain a minimum cement content. Otherwise, there is too little water in the mix to ensure proper compaction of the concrete. Proper measures should also be taken to prevent shrinkage cracking that can compromise the impermeability of the concrete surface.

Durability of concrete can be improved through the incorporation of polymers. Latex can be added to concrete at the fresh state. After drying, a continuous polymer film will form as a lining in the capillary pores and microcracks. The permeability of concrete is then significantly reduced. To improve the durability of hardened concrete in existing structures, polymer impregnation can be carried out. The structure is first heated to remove water in the capillary pores. Monomers of low viscosity (such as methyl methacrylate and styrene) are then allowed to penetrate into the concrete. Subsequent heating leads to polymerization and the formation of a highly impermeable layer on the surface of the structure. When high chemical resistance is required, cement can be completely replaced by polymers such as epoxy and polyester. Polymer concrete (which consists of polymer and aggregates alone) is very costly. Its application is often limited to repair jobs where durability and early strength are major concerns.

6. *High-performance Concrete and Other Developments*

High-performance concrete may be defined as concrete with strength and durability significantly beyond those obtained by normal means. The required properties for concrete to be classified as high performance therefore depend on the properties of normal concrete achievable at a particular time and location. At the present time, high-performance concrete in developed countries usually refers to concrete with 28-day compressive strength beyond 70–80 MPa, durability factor (defined as the percentage of original modulus retained after 300 freeze/thaw cycles) above 80%, and w/c below 0.35. It is made with good quality aggregates, high cement content (450–550 $\mathrm{kg\,m^{-3}}$), and a high dosage of both silica fume (5–15 wt.% of cement) and superplasticizer (5–15 $\mathrm{l\,m^{-3}}$). Sometimes other pozzolanic materials are also used. The high performance is achieved with the use of low w/c (0.20–0.35) as well as pozzolans to produce a dense microstructure that is high in strength and low in permeability. Superplasticizer is added to keep the mix workable.

With high cement content, the use of superpasticizers and silica fume and the need for more stringent quality control the unit cost of high-performance concrete can exceed that of normal concrete by 30–100%. Despite the higher material cost, the use of high-performance concrete is found to be economical for columns of tall buildings, as the amount of steel reinforcement can be reduced. In bridges, the reduction in deck size and weight effectively increases the allowable unsupported span. For a continuous bridge, the number of piers can be reduced. In many infrastructure projects, high-performance concrete is chosen for its durability against various types of chemical (e.g., sulfate or chloride) and physical attack (e.g., abrasion).

High-performance concrete can also be produced with lightweight aggregates. However, the aggregate needs to be very carefully chosen to make sure it is sufficiently strong. As long as the lightweight aggregate is strong enough, its use can indeed be advantageous. By saturating their pores with water before mixing, these aggregates can act as internal reservoirs that supply water to ensure continued cement hydration and prevent autogeneous shrinkage due to self-desiccation. This aspect is of particular relevance to concrete with a very low w/c, in which the early development of high density and low permeability makes it difficult for water to penetrate uniformly for the hydration process to continue.

Besides the production of high-performance concrete, superplasticizers are also commonly used in the production of high-workability concrete. With a slump value of 180–230 mm, high-workability concrete can be pumped rapidly over long distances, easily compacted in structures with highly congested reinforcement, and can even be self-compacting (i.e., requiring no external compaction effort). With superplasticizers, it is also possible to reduce the cement content while retaining the same workability. The possibility of thermal cracking in massive structures can therefore be reduced.

In the continual quest for improving concrete performance, it was soon realized that the size of aggregates is an important factor. By using very fine aggregates, superplasticizers, and a high dose of silica fume (about 20–30% of the cement content) concrete strength beyond 200 MPa can now be achieved by conventional techniques. One example is DSP—densified system with fine particles. Using strong aggregates of small size (e.g., calcined bauxite with maximum size of 4 mm), DSP with compressive strength over 250 MPa can be produced. Reaction powder concrete (RPC) is another example. With the maximum particle size limited to 0.4 mm, the compressive strength reaches 170 MPa by 28 days under room temperature curing. Curing at 80–90 °C will further increase the strength to 230 MPa. If pressure is applied before and during setting, and curing is carried out at 400 °C, strength as high as 680 MPa can be attained. With very high strength, both DSP and RPC are extremely brittle. Fiber reinforcement is therefore essential to prevent catastrophic failure at ultimate load.

Bibliography

ACI Committee 211.1 1999 *Standard Practice for Selecting Proportions for Normal, Heavyweight, and Mass Concrete.* ACI Manual of Concrete Practice, Part I

Aitcin P C 1998 *High-performance Concrete.* E. & F. N. Spon, London

Balaguru P N, Shah S P 1992 *Fiber-reinforced Cement Composites.* McGraw-Hill, New York

Bentur A, Diamond S, Berke N 1997 *Steel Corrosion in Concrete.* E. & F. N. Spon, London

Bentur A, Mindess S 1990 *Fibre Reinforced Cementitious Composites.* Elsevier Applied Science, London

Chandra S, Ohama Y 1994 *Polymers in Concrete.* CRC Press, Boca Raton, FL

Day K W 1999 *Concrete Mix Design, Quality Control and Specification*, 2nd edn. E. & F. N. Spon, New York

Elfgren L (ed.) 1989 *Fracture Mechanics of Concrete Structures, from Theory to Applications.* Chapman & Hall, London

Klieger P, Lamond J F (eds.) 1994 *Significance of Tests and Properties of Concrete and Concrete-making Materials.* ASTM STP 169C, ASTM Philadelphia

Malhotra V M, Carino N J (eds.) 1991 *Handbook on Nondestructive Testing of Concrete.* CRC Press, Boca Raton, FL

Mehta P K, Monteiro P J M 1993 *Concrete—Structure, Properties and Materials*, 2nd edn. Prentice Hall, Englewood Cliffs, NJ

Meitz J *Electrochemical Rehabilitation Methods for Reinforced Concrete Structures—A State of the Art Report.* IOM Communications, London

Mindess S, Young J F 1981 *Concrete.* Prentice Hall, Englewood Cliffs, NJ

Neville A M 1995 *Properties of Concrete*, 4th edn. Wiley, New York

Popovics S 1992 *Concrete Materials—Properties, Specifications and Testing.* Noyes Publications, NJ

Short A, Kinniburgh W 1978 *Lightweight Concrete.* Applied Science Publishers, Barking, UK

US Department of the Interior 1981 *Concrete Manual: a Manual for the Control of Concrete Construction.* US Department of the Interior, Water and Power Resources, Wiley, New York

C. K. Y. Leung

Hong Kong University of Science and Technology,

Kowloon, Hong Kong

Concrete: Failure Mechanics

Concrete is a heterogeneous engineering material made from hydraulic cement (most common is Portland cement), aggregates, and water. Modern concretes contain many different additives to improve the density, compressive strength, and workability of the material. It is used extensively in the building and civil engineering construction industry. For the design of sound economic and durable structures knowledge of fracture mechanics of concrete is essential. For structural design, fracture mechanics helps to assess the limit state of the structure under consideration and the safety against collapse can be assessed. For development of new concrete composites, fracture mechanics may help in the design of materials with the best deformation, strength, and fracture toughness characteristics. This article deals with the failure mechanisms of this composite material at the micro-, meso-, and macroscale.

1. *Size and Scale Levels*

In principle the size range can be extended, but for practical purposes these three levels suffice. At the microlevel the structure of cement is widely accepted. The characteristic size is 10^{-6} m. The buildup of the cement from calcium silicate hydrates (CSH), calcium hydroxide, and water films between the different microstructural elements determines the strength and fracture toughness. Furthermore, the porosity either in the cement structure itself (gel pores) or between the CSH crystals as open space left over after excess water has evaporated from the cement is important for strength and toughness. The structure of hydrated cement is shown schematically in the lower left of Fig. 1. Observing the cement with larger or smaller magnification reveals more or less fine detail. In fact the fluctuations in microstructure diminish when the scale increases. This is shown schematically in the upper part of Fig. 1.

With increasing size scale, the cement can be regarded as a continuum: the microstructure is small with regard to the size of the structural element. For concrete, the mesolevel is now reached. The characteristic size is 10^{-3} m, and individual aggregate particles, larger entrained air bubbles, patches of

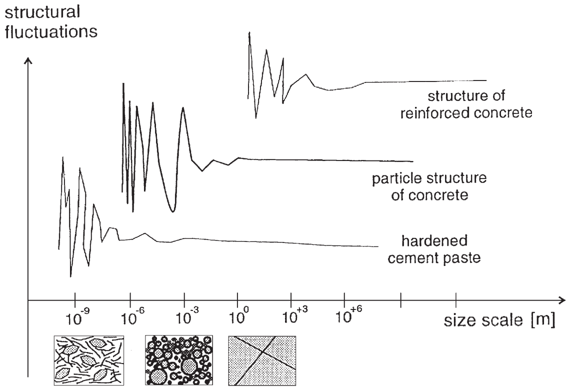


Figure 1
Size/scale levels for cement and (reinforced) concrete.

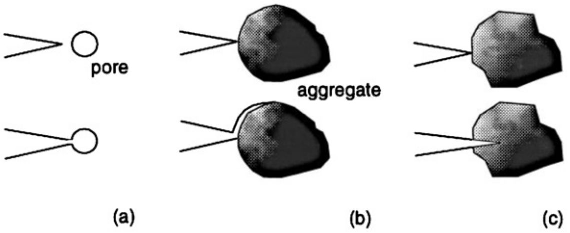


Figure 2
(a) Crack tip blunting. (b) Crack deflection in the vicinity of stiff and strong inclusions and (c) through cracking in weak inclusions.

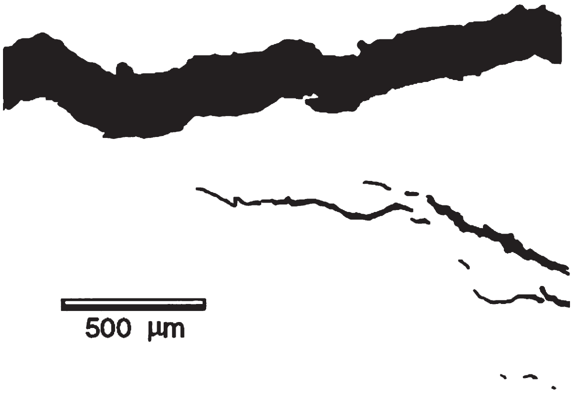


Figure 3
Multiple cracking leading to branching and bridging (after van Mier 1991).

hardened cement, and a distinct open interfacial transition zone between aggregates and cement matrix are observed. Again, the mesostructural fluctuations are large at a small scale, and diminish with decreasing resolution. The mesolevel with varying mesostructural fluctuations is shown schematically in the lower center of Fig. 1. One of the most important features at the mesolevel is the aggregate skeleton of the concrete.

Finally, the macrolevel is reached on further decreasing resolution. At this level the structural fluctuations from the cement microstructure and aggregate mesostructure cannot be observed any more, and the material is considered as a continuum (lower right of Fig. 1). Applications of concrete in structures are commonly in combination with steel reinforcement as the concrete itself has little bearing capacity in tension. At the macrolevel, the reinforcement forms its own structure, and again (macro-) structural fluctuations will appear when analyzing the behavior of reinforced concrete structures.

Modern concretes contain additions such as condensed silica fume to improve the uniaxial compressive strength, even above 200 MPa, the ductility, or fracture toughness with the addition of smaller or larger quantities of fibers. Attempts have also been made to improve the workability or the ease to place the material, which is achieved by additions such as superplasticizers and by choosing a well-balanced aggregate grading (Okamura *et al.* 1993). This latter type of material is called self-compacting concrete and is positioned without adding vibrational energy.

When subjected to loading, the material structure is prone to cracking. The applied loading can be mechanical, physical (temperature and moisture gradients), or chemical in nature, or combinations thereof; it can be uniaxial or multidirectional. Under low confinement, fracture of the material is generally quasibrittle. The cracks are almost immediately stress-free, although different types of mechanisms at different scales may enhance the relative ductility. The heterogeneity of the material acts at different size/scale levels both as an arrest mechanism and an accelerator mechanism.

2. *Crack Propagation*

At the micro- and mesolevel, small cracks may be arrested when they coalesce with a small pore. The mechanism can be regarded as crack tip blunting and is shown schematically in Fig. 2(a). Depending on the type of inclusion, cracks can be deflected (Fig. 2(b), high E-contrast and high strength of the inclusion) or propagate through the inclusion (Fig. 2(c), at low E-contrast and low strength).

Figure 3 shows most levels of crack growth in concrete subjected to uniaxial tension. Distributed cracking may occur in the tip zone of larger macroscopic cracks. Moreover intact regions in macroscopic cracks, often referred to as "crack face bridging," have been observed at many different size/scale levels (Fig. 4). One of the leading factors in the bridging mechanisms is the relative size of stiff aggregate particles, see Schlangen and van Mier (1992). The larger the maximum size of the stiff aggregates, the larger the bridging stress. The bridging of a crack is a three-dimensional phenomenon, and appears in many different materials at different size scales (van Mier 1997).

Next to aggregate bridging, load transfer may also occur through fibers bridging the crack. Again, this type of bridging mechanism appears to be governed by the fiber length (aspect ratio), the bond of the fiber to the cement matrix, and the fiber properties such as elasticity and strength.

3. *Models*

Heterogeneity of the material at different scales has an important influence on the specific fracture mechanisms. Once the mechanisms have been identified, modeling the fracture process becomes possible. Different approaches may be selected, namely a continuum approach where the different toughening mechanisms are implicit in the (mostly) nonlinear fracture law that is used, or it is possible to develop a material model in which the structure of the material is modeled explicitly. Depending on the level of operation, these models are referred to as "micromechanical" or "mesomechanical" models. An example of a nonlinear macroscopic model for concrete fracture is the "fictitious crack model," which was developed by Hillerborg and co-workers as an extension of the Dugdale–Barenblatt plastic crack tip model (Hillerborg *et al.* 1976). The main problem of the fictitious crack model is the parameter identification. Hillerborg proposed the use of a uniaxial tension test on a small double-notched concrete prism to measure the fracture energy and the shape of the softening curve in tension, as shown in Fig. 5. The fracture energy, tensile strength, and shape of the softening curve strongly depend on the size of the specimen and the adopted boundary conditions in the experiments. This implies that a unique determination of the parameters is almost impossible. Approximate use of the macroscopic models is possible only.

The micromechanical and mesomechanical models have the advantage that failure mechanisms are included directly. These models are often incorporated in a numerical model, such as those developed by Roelfstra *et al.* (1985) and Schlangen and van Mier (1992). The latter approach is based on developments in theoretical physics and starts from a regular or random triangular lattice of beam elements. A realistic structure of the material is projected on top of the lattice, and after the elastic properties are matched to those of a representative volume element of the real material, fracture studies can be performed. Removing the (in each computation step only one) beam element with the highest stress over strength ratio simulates fracture. In this way local elastic–purely brittle behavior is modeled. An example of an analysis with the model showing the effect of aggregate content on the fracture patterns under uniaxial tension is shown in Fig. 6.

When a high aggregate fraction is included, the bond zones around the aggregates touch and the percolation threshold for the bond zone is reached. In other words, continuous paths of connected bond beams can be found from one side of the structure to the other side. This implies that under uniaxial tension fracture is dominated by the behavior of the bond elements. Consequently relatively low structure strength is measured in this case (see Figs. 6(c) and (f)). The other extreme is when very sparse aggregates are included, as in Figs. 6(a) and (d). Now continuous paths of matrix beams connect the two specimen sides and the behavior is dominated by the matrix strength. Figure 6(d) shows that before the peak is attained a relatively long nonlinear portion is observed. The

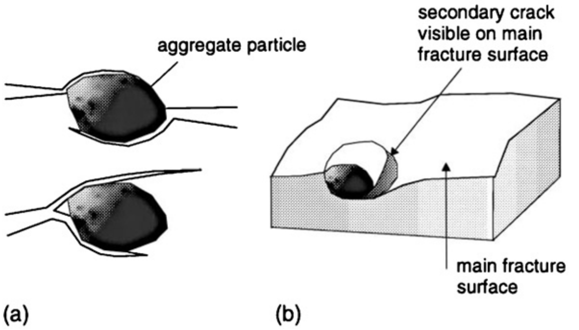


Figure 4
Crack face bridging mechanism: (a) two- and (b) three-dimensional view.

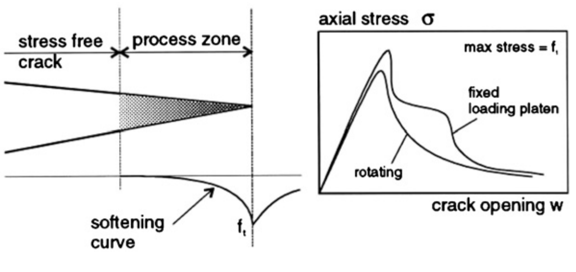


Figure 5
(a) Fictitious crack model and (b) softening stress crack opening diagram from a uniaxial tension test. The response in a tensile test clearly depends on the allowable rotations of the loading platens.

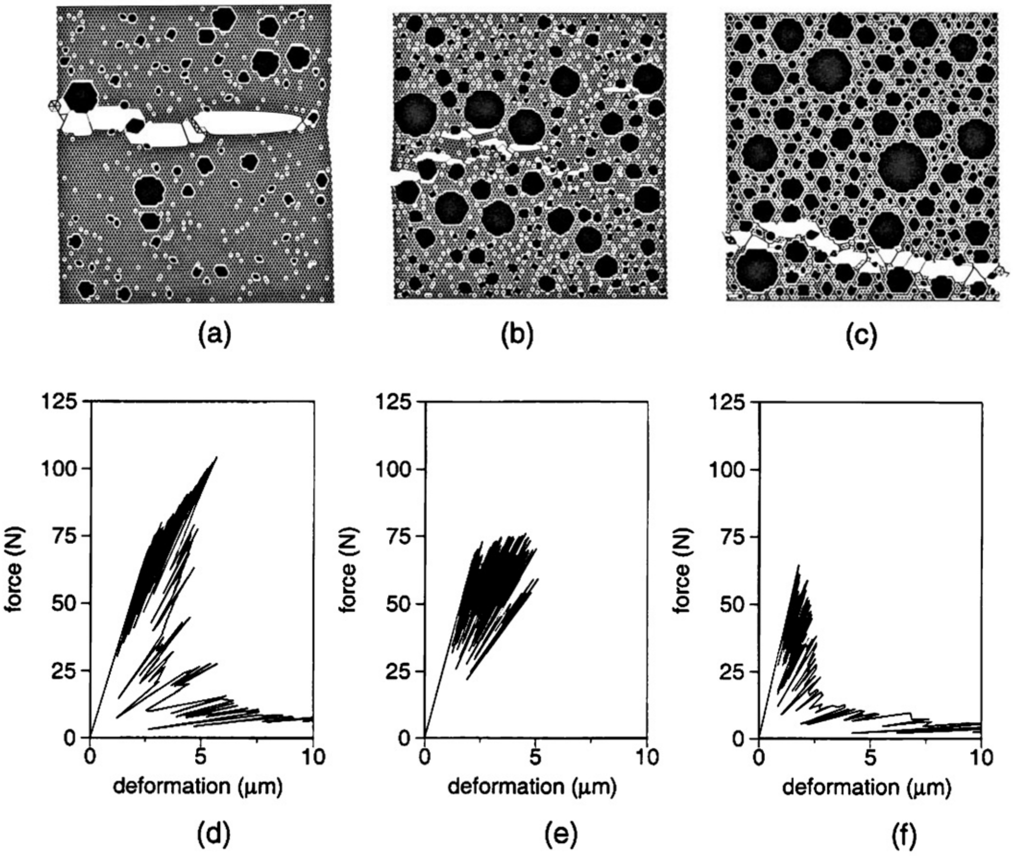


Figure 6
Example of lattice analysis of fracture of concrete under uniaxial tension. Different aggregate quantities are (a) 20%, (b) 50%, and (c) 80%. (d)–(f) show the corresponding load–deformation diagrams (after van Mier and Van Vliet 1999).

breaking of the weak bond elements around the aggregates causes this nonlinearity. In the examples shown in Fig. 6, the ratio of bond to matrix to aggregate strength is the most important parameter. The ratios are set at 1:4:8. Many of the toughening mechanisms mentioned at the beginning of this article can be recognized from the crack patterns of Figs. 6(a)–(c).

Another example is the simulation of a Brazilian splitting test (Lilliu *et al.* 1999), see Fig. 7. In this case strong boundary condition effects are measured. If the test is carried out in deformation control, i.e., by controlling the increase of the width of the vertical splitting crack, a secondary mechanism develops including radial bending cracks that grow from the edges of the cylinder to the main splitting crack.

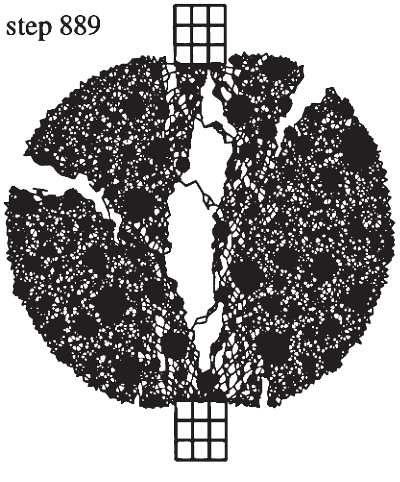


Figure 7
Simulated crack growth in a concrete disk subjected to a vertical splitting load (after Lilliu *et al.* 1999).

Bibliography

Hillerborg A, Modeer M, Petersson P-E 1976 Analysis of crack formation and crack growth in concrete by means of fracture mechanics and finite elements. *Cement Concrete Res.* **6**, 773–82

Lilliu G, van Mier J G M, Van Vliet M R A 1999 Analysis of crack growth of the Brazilian test: experiments and lattice analysis. In: Ellyin J, Provan J W (eds.) *Proc. ICM8 Progress in Mechanical Behaviour of Materials: Vol. I. Fatigue and Fracture*, pp. 273–8

Okamura H, Maekawa K, Ozawa K 1993 *High Performance Concrete*. Gihodo Shuppan, Tokyo (in Japanese)

Reinhardt H W, Naaman A E 1999 *High Performance Fibre Reinforced Cement Composites—3 (HPFRCC3)*. Rilem Publications, Paris

Richard P, Cherezy M 1994 Reactive powder concrete with high ductility and 200–800 MPa compressive strength. *Proc. ACI Spring Convention, SP144-24 (91AB)*. American Concrete Institute, Detroit, MI

Roelfstra P E, Sadouki H, Wittmann F H 1985 Le Béton Numérique. *Mater. Struct. (RILEM)* **26**, 327–35 (in French)

Schlangen E, van Mier J G M 1992 Experimental and numerical analysis of the micro-mechanisms of fracture of cement-based composites. *Cement Concrete Compos.* **14**, 105–18

van Mier J G M 1991 Crack face bridging in normal, high strength and lytag contact. In: van Mier J G M, Rots J G, Bakker A (eds.) *Fracture Processes in Concrete, Rock and Ceramics*. E & F N Spon, London, pp. 27–40

van Mier J G M 1997 *Fracture Processes of Concrete*. CRC Press, Boca Raton, FL

van Mier J G M, Van Vliet M R A 1999 Experimentation, numerical simulation and the role of engineering judgment in the fracture mechanics of concrete and concrete structures. *Constr. Build. Mater.* **13**, 3–14

J. G. M. van Mier
Delft University of Technology, The Netherlands

Construction: Cellular Materials

Many materials have a cellular structure made up of an interconnected network of struts or plates. Such cellular materials occur in nature as wood, cancellous bone, and cork. There are also many man-made cellular materials, both in the form of honeycomb-like materials and foams. Although initially honeycombs and foams could only be made from a limited range of materials, both can now be made from a wide range of polymers, metals, ceramics, and glasses. Micrographs of several natural and man-made cellular materials are shown in Fig. 1.

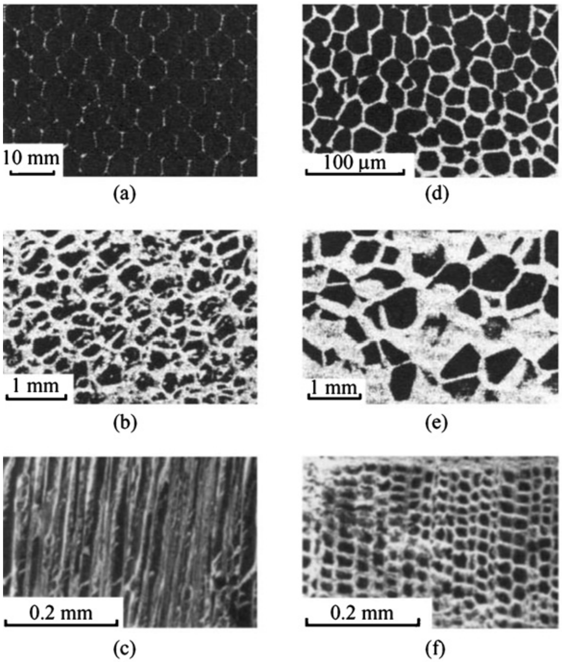


Figure 1
Micrographs of cellular materials: (a) aluminum honeycomb, (b) cork, (c) polyurethane foam, (d) polyethylene foam, (e) and (f) orthogonal views of cedar (reproduced by permission of Cambridge University Press from 'Cellular Solids: Structure and Properties', 1997).

Cellular materials currently are used extensively in a wide variety of applications. Wood artifacts have been found in Egypt dating back 5000 years. Approximately 1×10^9 t of wood are produced annually (Gibson and Ashby 1997). They are used extensively in lightweight sandwich panels. Separating two strong, stiff faces with a lightweight core creates a structure resistant to bending or buckling. They are a fine choice for insulation because of the low volume percentage of solids and their generally small cell size. Due to their low weight and ability to undergo large deformations at low loads, they are used by the packaging industry to protect materials from impacts.

This article summarizes the structure and mechanical behavior of cellular materials, production methods, and their uses in construction.

1. Structure and Mechanical Behavior

The properties of cellular materials are determined by a number of geometric and structural properties. The most important property is the relative density, determined by the density of the material divided by the density of the solid material from which it is made.

Cells generally extend in one of three forms: a 2D honeycomb, 3D open foam, or a 3D closed foam. The foam is determined to be open or closed based on whether or not the cell faces are covered by the solid material or empty. These structures are all shown in Fig. 1.

The cells are often anisotropic and often display anisotropic mechanical properties. Anisotropy is evident when the cells are elongated or have varying mean intercept lengths. This is often the case in rising

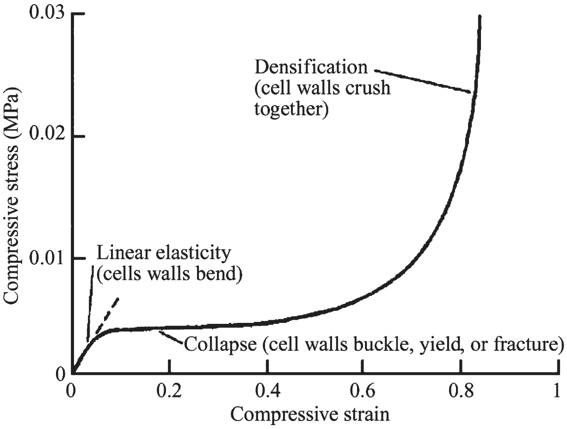


Figure 2
A typical stress–strain curve for a cellular material in uniaxial compression.

foams where cells elongate in one direction. Another common case is in wood, which accounts for some of the difference in behavior when loads are placed along or across the grain. Figure 1 shows this property.

A typical stress–strain curve for a cellular material is shown in Fig. 2. There are three distinct regimes of behavior. At relatively low loads, the material is linear elastic. At some critical level of load, the cells begin to collapse by elastic buckling, plastic yielding, or brittle fracture, depending on the nature of the cell wall properties. Cell collapse then progresses at a roughly constant load, producing a roughly horizontal stress plateau, until, at relatively large strains (typically about 0.8), the opposing cell walls begin to meet and touch. The stress then rises steeply as the material densifies (Maiti *et al.* 1984).

Each of these regimes of behavior is related to the mechanism by which the cells deform within that regime. Observation of model honeycomb materials and of foams in a scanning electron microscope reveals that the initial linear elasticity is produced by bending in the cell walls. Cell collapse is caused by elastic buckling in elastomeric honeycombs and foams, by plastic yielding in materials which yield, and by brittle crushing in materials which fracture (Gibson and Ashby 1982, Maiti *et al.* 1984). The strain at which the honeycomb or foam begins to densify depends on its relative density.

The behavior of foams in tension is similar, but there are some notable differences. The linear elastic period is determined by the same modulus of the material in compression. At slightly larger strains buckling as in compression is not possible. Instead cell edges rotate toward the tensile axis. After a strain of approximately 0.3 increasing strain requires the stretching of the cell walls. These stages are demonstrated in Fig. 3. Failure in tension is by the propagation of a single crack.

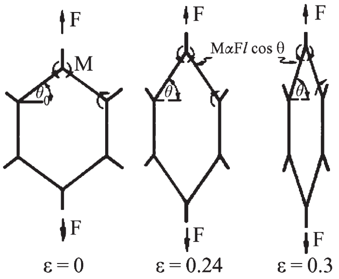


Figure 3
Alignment of the cell edges during tensile loading. The limiting strain is about 0.3 (reproduced by permission of Cambridge University Press from ‘Cellular Solids: Structure and Properties’, 1997).

This stress–strain behavior can be characterized by several mechanical properties. The linear elastic regime is described by a set of elastic moduli such as the Young’s modulus E, the Poisson ratio ν, and the shear modulus G. The plateau stresses corresponding to elastic buckling, plastic yielding, and brittle crushing are σ_{el}, σ_{pl}, and σ_{cr}, the limiting strain for densification is ε_D, and the tensile fracture toughness is K_{Ic}. By modeling the structure of cellular materials and the mechanisms by which they deform, relationships between the properties of the cellular material and its relative density, its solid cell wall properties, and its geometry can be found. Table 1 summarizes these relationships for both honeycombs and foams. A more detailed analysis can be found in Maiti *et al.* (1984).

2. *Production Methods*

A wide variety of materials can be used to create cellular solids. Production methods depend upon the structure and material of the cellular solid being formed.

Honeycomb structures are commonly made in four ways. They can be pressed into half hexagonal shapes and then glued together. Alternatively, glue can be applied in strips to flat sheets, stacked, and then expanded to create the honeycomb shape. Honeycombs can also be created by casting into a mold or by extrusion.

Foams are made in a variety of ways depending upon the material being used. Polymers are generally foamed by introducing gas bubbles into a hot polymer using either mechanical stirring or by mixing a blowing agent into the polymer. Glass foams are made using similar methods to polymers. Metallic foams are made using either liquid or solid state processing (Shapovalov 1994, Davies and Zhen 1983).

Table 1
Property-density relationships for cellular materials.

		Honeycombs	
Property	Open celled foams	In-plane	Axial
E	$C(\rho/\rho_{\mathrm{s}})^2E_{\mathrm{s}}$	$C(\rho/\rho_{\mathrm{s}})^3E_{\mathrm{s}}$	$C(\rho/\rho_{\mathrm{s}})E_{\mathrm{s}}$
v	C	C	C
G	$C(\rho/\rho_{\mathrm{s}})^2E_{\mathrm{s}}$	$C(\rho/\rho_{\mathrm{s}})^3E_{\mathrm{s}}$	$C(\rho/\rho_{\mathrm{s}})E_{\mathrm{s}}$
σ_{el}	$C(\rho/\rho_{\mathrm{s}})^2E_{\mathrm{s}}$	$C(\rho/\rho_{\mathrm{s}})^3E_{\mathrm{s}}$	$C(\rho/\rho_{\mathrm{s}})^3E_{\mathrm{s}}$
σ_{pl}	$C(\rho/\rho_{\mathrm{s}})^{3/2}\sigma_{\mathrm{ys}}$	$C(\rho/\rho_{\mathrm{s}})^2\sigma_{\mathrm{ys}}$	$C(\rho/\rho_{\mathrm{s}})^2\sigma_{\mathrm{ys}}$
σ_{cr}	$C(\rho/\rho_{\mathrm{s}})^{3/2}\sigma_{\mathrm{fs}}$	$C(\rho/\rho_{\mathrm{s}})^2\sigma_{\mathrm{fs}}$	$C(\rho/\rho_{\mathrm{s}})\sigma_{\mathrm{fs}}$
K_{Ic}	$C(\rho/\rho_s)^{3/2}\sigma_{fs}\sqrt{\pi l}$	$C(\rho/\rho_s)^2\sigma_{fs}\sqrt{\pi l >}$	$C(\rho/\rho_s)\sigma_{fs}\sqrt{\pi l}$

The constants of proportionality are different for each property and each case (foams, honeycomb-in-plane, and honeycomb-axial). E_{s}, σ_{ys}, and σ_{fs} are the Young's modulus, yield strength and modulus of rupture, respectively, of the solid cell-wall material. The characteristic length of the cell size is l.

3. *Applications of Cellular Materials in Construction*

3.1 *Sandwich Panels*

Structural panels made up of two stiff, strong faces separated by a lightweight core are known as *sandwich panels*. The core acts to separate the faces, increasing the moment of inertia of the panel, with little increase in weight, producing an efficient structure for resisting bending and buckling loads. Usually the core is made up of a honeycomb, end-grain balsa wood, or cellular foams made from metals, polymers, or glass. The high stiffness and strength to weight ratios make sandwich panel construction successful in their use in space craft, aircraft structures, marine vessels, train and truck structures, containers, tanks, car body parts, and building structures (Burman and Zenkert 1997).

The technology for producing sandwich panels was first developed in the aerospace industry, beginning with the balsa core-plywood face sandwiches of the World War II Mosquito aircraft (Allen 1969). Modern aerospace sandwich components are now made almost exclusively with honeycomb cores and fiber-composite faces bonded together with advanced adhesives. Sandwich panels are now made routinely and are reliable for components such as the rotor blades of helicopters, wing tip sections, tail sections, and flooring panels. This technology is now well developed and is being transferred to the construction industry.

The main use of sandwich panels in the construction industry is in building panels. Sandwich construction is used in wall, roof, and door paneling. Different combinations of face and core materials are used; typically the faces are made of steel, aluminum, or waferboard and the cores are made of foamed polyurethane, foamed polystyrene beadboard, or foamed glass.

Sandwich panels provide a strong, yet economic, material for construction. Their rigid faces provide strength and stiffness to a structure while their lightweight centers often provide insulative as well as important construction properties.

The lightweight, yet rigid, property of sandwich panels allows for other current applications. Sandwich design is used in the development of downhill skis, domed covers of water reservoirs, and even the folding bridges employed by the US Army.

3.2 *Wood*

Wood is still one of the most widely used structural materials throughout the world.

The cellular structure of wood is shown in Fig. 1. Softwoods used in construction have two main types of cells: tracheids and parenchyma. The tracheids are the long, narrow cells (aspect ratio 100:1) which make up the bulk of the tree. They conduct fluids and nutrients within the tree and give it its structural support. The parenchyma cells, found in the rays, store sugars.

At a finer level, the walls of cells have a fiber-composite structure with partially crystallized cellulose molecules forming stiff, strong, fiber-like chains helically wound in a matrix of amorphous lignin and semicrystalline hemicellulose. Each cell wall is made up of four layers, each with a different orientation of cellulose fibers.

The mechanical behavior of wood can be modeled by analyzing the mechanisms by which it deforms (Easterling *et al.* 1982). When loaded across the grain, the cell walls bend in the same way that a honeycomb does when loaded in the plane of the honeycomb. Because of this, the Young's moduli and compressive strength of woods loaded across the grain depend on the cube and the square of the relative density of the wood. Figures 4 and 5 show this. When loaded along the grain, the cell walls compress uniaxially, as the honeycomb does if loaded normal to the plane of the hexagonal cells. Because of this the

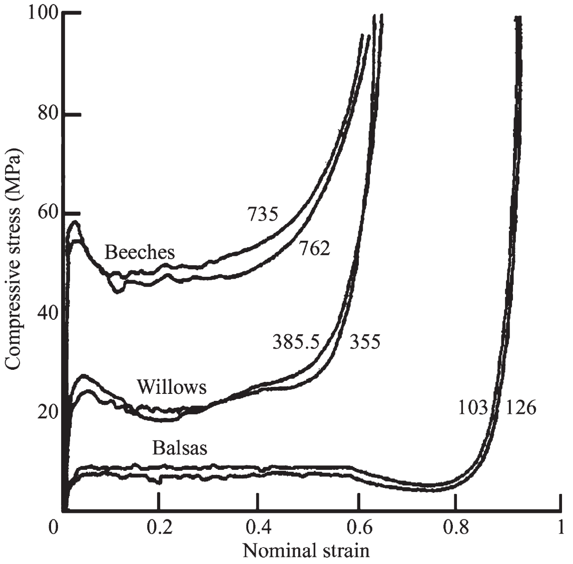


Figure 4
Stress–strain curves for wood loaded along the grain. Figures by curves indicate density (kg m^{-3}); axial compression = 1.25–1.50 × 10^{-5} s^{-1} (reproduced by permission of Cambridge University Press from 'Cellular Solids: Structure and Properties', 1997).

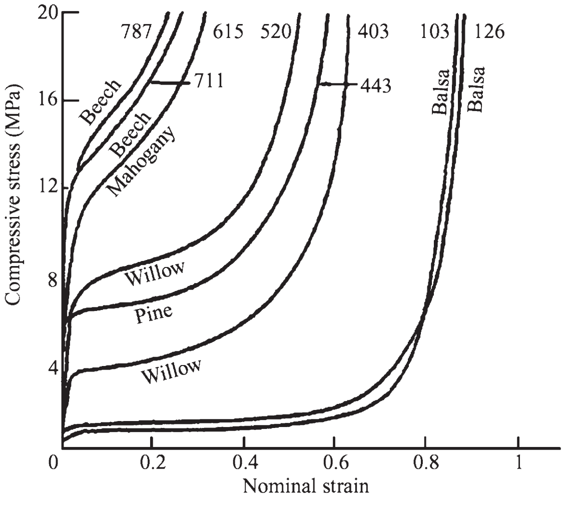


Figure 5
Stress–strain curves for wood loaded across the grain. Figures by curves indicate density (kg m^{-3}); tangential compression = 1.6 × 10^{-3} s^{-1} (reproduced by permission of Cambridge University Press from 'Cellular Solids: Structure and Properties', 1997).

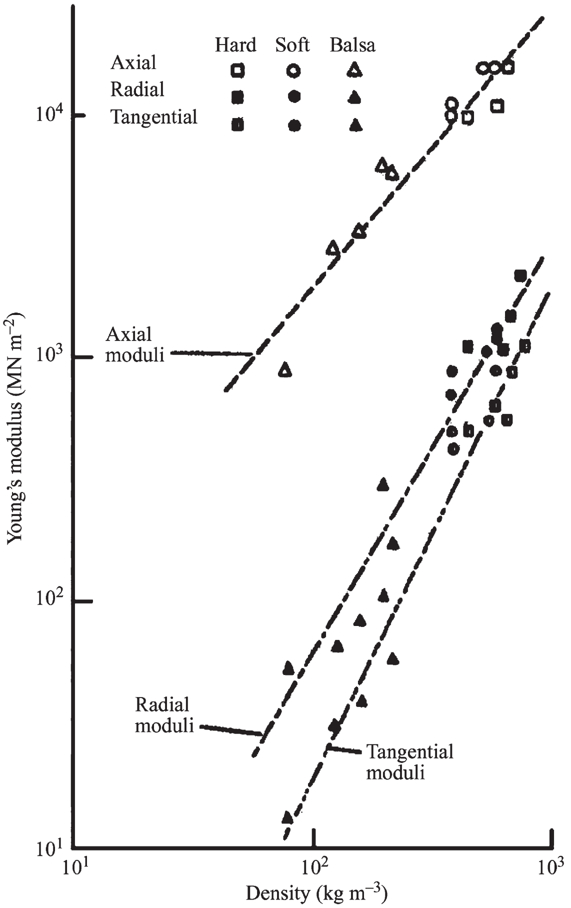


Figure 6
Young's modulus for wood, plotted against density. The lines indicate the theoretical dependencies (reproduced by permission of The Royal Society from *Proc. Roy. Soc. A*, 1982, **383**, 31–41).

Young's modulus along the grain depends linearly on the relative density of wood. Finally, the axial compressive strength of wood is related to yielding in the cell walls, again giving a linear dependence on relative density. These results are plotted in Figs. 6 and 7, along with data for wood. The modeling begins to explain the dependence of the properties of wood on the loading direction and the relative density of the wood. The tensile fracture toughness of wood has been explained in a similar way; see for example, Ashby *et al.* (1985).

The mechanical properties of wood tend to be variable, due to the presence of defects in the wood, such as nonparallel grain and knots. Recently, glue-laminated wood members (glulam) have been developed to minimize the effect of such defects. Gluing together thin strips of wood, from which the defects

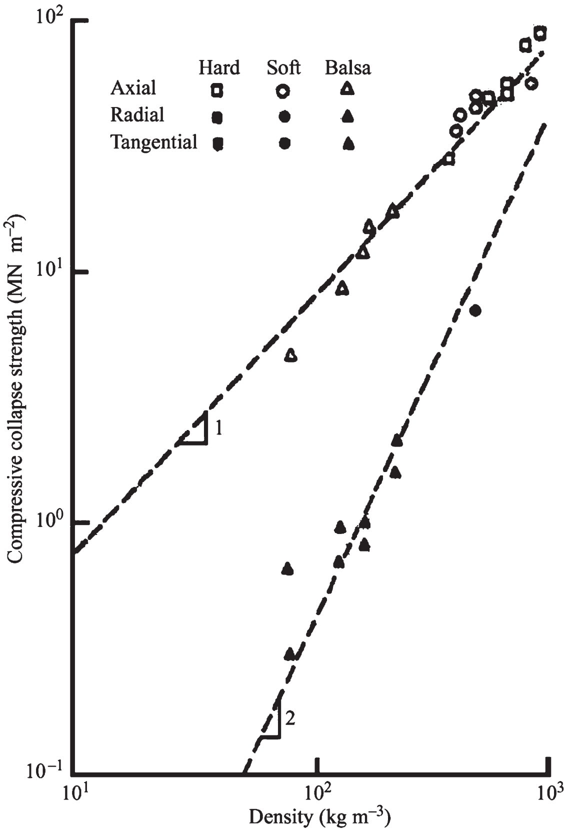


Figure 7
Compressive strength for wood, plotted against density. The lines indicate the theoretical dependencies (reproduced by permission of The Royal Society from *Proc. Roy. Soc. A*, 1982, **383**, 31–41).

have been cut out, makes glue-laminated members. The properties of such members are more uniform than those of dimensioned lumber. Glulam sections also have the advantage that they can be curved and built up to any size. The efficiency of wood members can also be improved by using them in sandwich constructions. Typically, plywood faces are used in conjunction with honeycomb, foam, or balsa cores.

3.3 *Foams and Polymers*

Cellular forms of thermal, electrical, and acoustic insulation exist as foams and polymers made of metals, glass, or plastics (polyethylene, polystyrene, polyurethane, etc.). Foam and polymer insulation have many properties allowing them to perform well. Foams have low thermal conductivity (perfect for thermal insulation) (see Table 2), low dielectric loss (allowing microwave transmission without attenuation or scattering), and sound absorbency capabilities (providing great acoustical insulating properties) (Gibson and Ashby 1997).

Table 2
Thermal conductivities and diffusiveness.

Material	Thermal conductivity, l (Wm^{-1}K^{-1})	Thermal diffusivity, α (m^{2}s^{-1})
Glass (solid)	1.1	8.9×10^{-5}
Polyurethane (solid)	0.25	
Polystyrene (solid)	0.15	1.0×10^{-7}
Air	0.025	
CO_2	0.016	
Balsa ($\rho^*/\rho_s=0.09$)	0.055	
Cork ($\rho^*/\rho_s=0.14$)	0.045	
Polystyrene foam ($\rho^*/\rho_s=0.025$)	0.040	1.1×10^{-6}
Polyurethane foam ($\rho^*/\rho_s=0.02$)	0.025	9.0×10^{-7}
Glass foam ($\rho^*/\rho_s=0.05$)	0.050	

Bibliography

Ackers P 1945 The efficiency of sandwich struts utilizing calcium alginate core, R & M 2015. Aeronautical Research Council, Farnborough, UK

Allen H G 1969 *Analysis and Design of Structural Sandwich Panels*. Pergamon, Oxford

Ashby M F, Easterling K E, Harrysson R 1985 The fracture toughness of woods. *Proc. Roy. Soc. A* **398**, 261–80

Bodig J, Jayne B A 1982 *Mechanics of Wood and Wood Composites*. Van Nostrand Reinhold, New York

Burman M, Zenkert D 1997 Fatigue of foam core sandwich beams-1: undamaged specimens. *Int. J. Fatigue* **19**, 551–61

Davies G J, Zhen S 1983 Metallic foams: their production, properties and applications. *J. Mater. Sci.* **18**, 1899–911

Demsetz L A, Gibson L J 1987 Minimum weight design for stiffness in sandwich plates with rigid foam cores. *Mater. Sci. Eng.* **85**, 33–42

Dinwoodie J M 1981 *Timber: Its Nature and Behaviour*. Van Nostrand Reinhold, New York

Easterling K E, Harrysson R, Gibson L J, Ashby M F 1982 On the mechanics of balsa and other woods. *Proc. Roy. Soc. A* **383**, 31–41

Gibson L J 1984 Optimization of stiffness in sandwich beams with rigid foam cores. *Mater. Sci. Eng.* **67**, 125–35

Gibson L J, Ashby M F 1982 The mechanics of three dimensional cellular materials. *Proc. Roy. Soc. A* **382**, 43–59

Gibson L J, Ashby M F 1997 *Cellular Solids: Structure and Properties*. Cambridge University Press, Cambridge

Glicksman L R 1994 Heat Transfer in Foams. In: Hilyard N C, Cunningham A (eds.) *Low Density Cellular Plastics*. Chapman & Hall, London Chap. 5

Handbook of Chemistry and Physics, 66th edn. 1985–1986 Chemical Rubber Co, Cleveland, OH

Huang S N, Alspaugh D W 1974 Minimum weight sandwich beam design. *AIAA J.* **12**, 1617–8

Kuenzi E W 1965 Minimum weight structural sandwich. US Forest Service, Research Note FPL-086. Forest Products Laboratory, Madison, WI

Maiti S K, Gibson L J, Ashby M F 1984 Deformation and energy absorption diagrams for cellular solids. *Acta Metall.* **32**, 1963–75

Patten G A, Skochdopole R E 1962 Environmental factors in thermal conditioning of plastic foams. *Mod. Plast.* **39**, 149

Schuetz M A, Glicksman L R 1983 *Proc. SPI 6th Int. Technical/Marketing Conf.* San Diego, CA, 332–40

Shapovalov V 1994 Porous metals. *Mater. Res. Bull.* **19**, 24–8

Ueng C E S, Liu T L 1979 Least weight of a sandwich panel In: Craig R R (ed.) 1979 *Proc. ASCE Engineering Mechanics Division 3rd Specialty Conf.* University of Texas, Austin, TX, pp. 41–4

Wittrick W H 1945 A theoretical analysis of the efficiency of sandwich construction under compressive end load, R & M 2016. Aeronautical Research Council, Farnborough, UK

F.-J. Ulm
Massachusetts Institute of Technology, Cambridge, Massachusetts, USA

Construction Materials: Lightweight Aggregates

Many materials serve as lightweight aggregates in concrete, plaster blocks, and formed products. Many are mineral raw materials, others are waste or synthetic products, and some are organic materials (Table 1).

The distinction between normal weight and lightweight aggregates is at about 75 lbft^{-3} (1200 kg m^{-3}) for loose graded aggregate. Most lightweight aggregates used for structural purposes are in the range 35–60 lb ft^{-3} (560–960 kg m^{-3}), those used for nonload-bearing purposes (ultralightweight aggregates) are in the range 5–20 lb ft^{-3} (80–320 kg m^{-3}).

Expanded clays, shales, and slates make up about 51% by weight of the lightweight aggregates used annually in the USA, with the remainder divided among scoria (see below), expanded slag, fly ash, vermiculite, and perlite. Vermiculite and perlite are the most used ultralightweight materials, but less than 25% is used as construction aggregate. They are used in about equal amounts. Pumice–pumicite and diatomite, in about a 10:1 ratio, constitute the remaining third of that market.

About 7.8 million short tons (7.1×10^6 t) of rock were used in 1997 for structural lightweight aggregates, and perhaps as much as 1.5 million short tons (1.36×10^6 t) were used for ultralightweight aggregates for nonload-bearing concretes, blocks, plaster insulation board, fillers, filters, agricultural aids, and loose fill insulation.

Excellent detailed review articles include Mason (1994) for structural lightweight aggregates, Geitgey (1994) for pumice and scoria, Breese and Barker (1994) for perlite, and Hindman (1994) for vermiculite.

1. Structural Lightweight Aggregates

Aggregates suitable for structural uses have several common ideal characteristics, high porosity, poor permeability, and relatively thick pore walls (compared to void diameter), that give the aggregates high compressive strengths. The porosity of all these aggregates results from heat treatment. A porous but relatively impermeable texture makes lightweight aggregates both thermal and acoustic insulators, desirable for bulk use in floor and roof fill and with an added advantage in saving weight in concrete.

Higher strength and increased weight go together. Compressive strength ranges, for example, from about 1200 psi (8.3×10^6 Pa) for 60 lb ft^{-3} (960 kg m^{-3}) scoria aggregate concrete to more than 6000 psi (42×10^6 Pa) for 120 lb ft^{-3} (1920 kg m^{-3}) expanded (sintered) shale concrete. In this broad range, lightweight aggregate concretes can be used for fill, bridge decks, shell walls, concrete floors, concrete frames using precast, prestressed, or post-tensioned girders, precast tilt-up units, poured barge and ship hulls, and lightweight masonry blocks. In bulk form, the aggregates are used as railroad ballast, as highway and embankment fill, and in the surface course of bituminous paving mixtures.

Structural lightweight aggregates allow higher buildings, more fire-resistant steel frameworks, and better insulated floors and ceilings. They save money, materials, and time because concrete can be pumped higher and further; plates, girders, and shell walls can be mass produced offsite; and additional thermal and acoustic insulation can be reduced. Additionally, lightweight aggregates and concretes allow innovative, experimental construction and design.

The USA has abundant domestic supplies; there are virtually no import or export markets. The materials have low unit value but high place value (transportation costs largely govern whether the materials can be used). Production data for pumice reveal some imports are used for lightweight block, but in 1997 about 0.29 million short tons (2.63×10^5 t) were imported for *all* purposes, as compared with nearly 1.5 millionshort tons (1.36×10^6 t) of domestic lightweight aggregates used in lightweight block. Almost all pumice imports (less than 2% of the market) are used along the eastern seaboard, where marine freight rates allow competition from domestic western resources.

1.1 Clays, Shales, and Slates

Many clays, shales, and slates expand somewhat on heating to about 2000 °F (1100 °C). The rocks differ in their physical characteristics, but are chemically,

Table 1
Classification and properties of lightweight aggregate materials.

Material	Composition (wt.%)	Texture (unexpanded)	Expansion needed	Texture (expanded)	Weight (lb ft^{-3}/kg m^{-3})	Remarks
Products of explosive volcanism						
Pumice	Silicic	Frothy, very vesicular	None	Unchanged	30–60/480–960	Blocks and fragments of coherent glass
Pumicite	Silicic	Porous	None	Unchanged	60–70/960–1120	Shards of glass; very fine grained; very low strength; expansible
Scoria	Basic	Clinker-like, very vesicular	None	Unchanged	35–75/560–1200	Tephra as bombs and lapilli; vesiculated zones in lava flows
Perlite	Silicic (H_2O:2–5)	Fractured, onionskin, partings	To × 6, ~1400 °F	Cellular, frothy spheres	6–16/100–250	Low strength; high permeability-to-surface area ratio; ultralight
Obsidian	Silicic (H_2O: <2)	Dense, glass, conchoidal fracture	To about × 5, 1400–1500 °F	Cellular, frothy spheres	25–30/400–480	Low strength; high permeability-to-surface area ratio; ultralight
Deposits of sedimentary origin						
Clays, shales, slates	Wide range; silty, calcareous, carbonaceous, ferruginous	Dense, may be fissile	To × 2–3, ~1800– 2200 °F	Cellular, fairly uniform	30–60/480–960	Kiln or sinter product; good strength; red ores least likely to expand
Diatomite	SiO_2: ~100; may be "clayey", petroliferous	Porous, may be friable	None	Cellular	30–50/480–800	Coproduct of calcining for filter-aid products
Alteration of mafic silicate minerals						
Vermiculite (hydrobiotite)	Hydrated Mg–Fe–Al silicates; varies considerably	Micaceous	To × 8–15, ~2000 °F	Accordian- like	6–15/100–240	Supergene alteration of biotite, phlogopite, hornblende, diopside, augite, olivine, chlorite, serpentine; low strength; ultralight

By-products of manufacturing processes						
Blast furnace	Best slag is SiO_2: 35–88; $CaO + Al_2O_3$: 44–47; low MgO, FeO, S	Liquid slag	About ×8, ambient temperature	Cellular, fairly	35–60/560–960	Zinc and phosphate smelter slags have also been used successfully
Fly ash, bottom ash, coke breeze	Wide ranges; $SiO_2 + Al_2O_3 + Fe_2O_3$: >70; MgO: <5; SO_2: <3; alkalis: $<1 \cdot 5$	Extremely fine grained	None	Cellular	35–55/560–880	Ash from burning pulverized coal; sintered to cause agglomeration
Cinders	Wide ranges; $SiO_2 + Al_2O_3 + Fe_2O_3$: >70; MgO: <5; SO_2: <3; alkalis: $<1 \cdot 5$	Cellular, nonuniform	None	Unchanged	40–60/640–960	From burning pea-sized or larger coal, coke
Synthetic and miscellaneous materials						
Plastic beads and foams	Polyurethane, polystyrene, polyvinyl chloride	Does not apply	Various temperatures, reactions	Cellular, uniform spheres	$<15/<240$	Many patented techniques and applications
Foaming compounds	Al powder, vinyl resins, H_2O_2	Does not apply	Various temperatures, reactions	Does not apply	Various	Does not substitute for aggregate; lowers density by adding voids
Rice hulls, wheat hulls, sawdust, wood fibers	Organic	Fibrous, cellular	None	Unchanged	No data	Generally low strength; very minor usage

Source: Bush (1973).

and to some extent mineralogically, similar. Their chemical composition makes them amenable to thermal expansion. Representative chemical analyses for argillaceous rocks suitable for use as lightweight aggregate are given in Bush (1973), p. 338). The range in composition is wide and overlaps for all three materials. Unfortunately, the analyses of rocks that expand suitably, inadequately, or not at all, also fall in these ranges.

The basic requirement for expansion is a sufficient source of gas within the raw material to cause a full bloat, in the pyroplastic temperature range, so that the gas can be trapped; there must be enough vitrification, at a high enough viscosity, so that the gas can be held in small, evenly distributed pores; and the temperature range between softening and liquefaction must be large enough (about 100 °F (40 °C)) that the bloating can be controlled in large-scale commercial production. Almost all production is in horizontal kilns at temperatures between 1800 and 2200 °F (980–1200 °C), although temperatures may be as low as 1600 °F (870 °C) or as high as 2400 °F (1300 °C).

There is little agreement on the relative importance of the sources for the gas evolved within the raw material. Details of numerous laboratory studies can be found in the summary article by Bush (1973). Conventional wisdom in the early days of the industry held that the argillaceous rocks required a significant organic content, but laboratory investigation of the evolved gases showed that many materials with very little total carbon expanded better than those with an appreciable amount, and that CO_2, SO_2, and H_2O all played a part in the bloating.

Additional studies suggested sources of CO_2, SO_2, and H_2O from thermal destruction of hydrous iron oxides (goethite, limonite, hydrated hematite), pyrite, calcite, dolomite, ankerite, and from sulfides and sulfates, from dehydroxylation of amphiboles and micas, and gas from chemical interactions between organic matter, iron oxides, and water. Perhaps some reactions may operate simultaneously.

A range of compositions (52–80% SiO_2, 11–25% Al_2O_3) and combined fluxes (10–25% Fe_2O_3, FeO, CaO, MgO, and $(K,Na)_2O$) that included most of the bloatable materials and excluded the nonbloatables was defined by Riley (1951) on the basis of 85 chemical analyses of clays and shales. Additional investigations since 1951 redefined this bloating range as White (1960) (with an additional 93 samples) showed that the range had to be extended to higher Al_2O_3 and lower flux content, whereas Sweeney and Hamlin (1965) extended the range to a higher flux content. Laboratory methods for evaluating raw materials are cited in Bush (1973). The ultimate evaluation is how several tons of material expands in a commercial operation.

Distinguishing favorable from unfavorable argillaceous rocks in the field is difficult because so many types are suitable. Illitic and montmorillonitic rocks and sediments seem to be more consistently favorable than kaolinitic ones. Dark-colored clays, shales, and slates are more suitable than light-colored (particularly red) ones, especially if they contain dispersed organic carbon. In general, unweathered materials are more suitable than weathered ones.

Common or brick clay is more suitable both physically and economically than flint or ball clay. Thinly fissile shales and slates tend to expand unidimensionally, which is undesirable. Most of these features tend to vary more in short vertical intervals across the rock units than laterally along them, but lateral variation is enough that units unsuitable at one point may be suitable a few hundreds yards to a few miles away.

Marine, littoral, lacustrine, and fluviatile clays, and shales and slates derived from them, are all possible raw materials. Geologic mapping in the USA, at scales adequate to identify the presence of argillaceous rocks, is sufficiently complete to virtually rule out the possibility of finding significant additional units.

In addition to the expandable materials, argillaceous rocks can be ground, mixed with combustible materials, such as ground coal, and sintered to produce lightweight aggregates. The mass becomes pyroplastic and voids form as the combustibles are burned out. Inasmuch as units that may not bloat of themselves can thus be used for lightweight aggregate, classification of materials as unfit is unwise unless testing has shown unsuitability for both processes.

Reserves of expandable clays, shales, and slates are so large that they will satisfy the USA's needs for many years.

1.2 *Volcanic Rocks*

Various volcanic rocks worldwide are suitable for use as lightweight aggregates; all are extrusive ejecta that range in composition from mafic to silicic. Scoria and volcanic cinder (both mafic), and pumice and pumicite (silicic) make moderate- to low-strength structural aggregates; expanded perlite and expanded pumicite are ultralightweight materials. Water content in basalt is little different from that in scoria and cinder, but the difference between water in rhyolite and that in silicic glassy rocks is significant. An excess of water allows the perlites and pumicites to be thermally expanded. In a general way, viscosity of extruded rock melt varies directly with the SiO_2, Na_2O, and K_2O content and inversely with the Fe_2O_3, FeO, CaO, and MgO content; the mafic rocks are more fluid and the silicic rocks are more likely to retain gases.

1.3 *Scoria and Volcanic Cinder*

Scoria (larger than 1 in (2.5 cm) in size) and volcanic cinder (smaller) are cellular, frothy products of explosive mafic volcanism—erupted as bombs, blocks, lapilli, volcanic gravel, or lava flows. Much of the material is lava forcefully ejected from the vent, vesiculating as volcanic gases expand and are entrapped by rapid cooling, either while airborne as ejecta or at the surface of flows; broken blocks and irregular fragments result from continued movement. There is a wide range in vesicle size (particularly in the ejecta) and in the thickness of the cell walls, resulting in a considerable variation in the density of the rocks. Scoria and cinder form cones and blankets near the vent; lava flows are fluid, extending laterally to tens of miles, and to tens of feet in thickness. Scoriaceous zones in flows are highly variable and are confined to the upper few feet.

Most of the deposits of scoria and volcanic cinder are geologically youthful, surface or near-surface deposits, some only a few thousand years old. Although ejecta enter readily into the sedimentary cycle, they lack sufficient resistance to abrasion to survive reworking for very long.

In the USA deposits of scoria and cinder are restricted to states west of the 104th meridian; production was recorded from all but Alaska. Arizona, Oregon, California, Hawaii, and New Mexico are the largest producers, generally in that order; however, in 1997 New Mexico was first. In the states west of the 104th meridian nearly all large cities (at least 50 000 people) and most cities of at least 25 000 are within about 100 miles (160 km) of economic deposits of scoria or cinder.

Available USA resources are estimated to be about 3 cubic miles (12 km^3) of scoria (based in part on Higgs (1949), about 9000 million short tons (8.2×10^9 t)). At a projected level of 3 million short tons (2.7×10^6 t) annually, scoria production for more than 1000 years can be adequately accommodated within this total resource.

1.4 Pumice and Pumicite

Pumice and pumicite are used both as structural lightweight aggregates and as ultralightweight aggregates. They are lighter colored, of lower density, and more highly cellular than are scoria and cinder. Composition ranges from high-silica (rhyolitic) to low-silica (basaltic), but most pumices are in the rhyolite range. High-silica rock melts generally lead to violent extrusion; gas and steam during explosive extrusion commonly shatter the pumice froth, and the shards and finer fragments form pumicite (volcanic ash). Low-density pumice and pumicite are thrown much further from the vents than scoria and cinder; many pumicite deposits are several hundred miles from their sources. In pumice, cell walls are commonly very thin, and the rock is light enough to float.

Deposits of pumice and pumicite commonly occur as an air-fall blanket, semiconcentric to a crater source vent or broadly fan-shaped lateral to fissure vents, sometimes with cones. Very large volumes occur in ash-flow tuffs and in the upper parts of rhyolite and "obsidian" domes. They are subject to rapid erosion, reworking, and redeposition as stream or lake sediments. Thicknesses range up to tens of feet, and areal extents range up to tens of square miles; amounts may be as large as many hundreds of thousands of tons.

Ash falls are air-sorted. Reworked and redeposited materials commonly are interbedded. Some ash falls also retain enough heat to weld and to form glassy zones. Most ash flows are unsorted or crudely sorted heterogeneous mixtures of pumice, ash, and rock fragments. The pumice as fragments and blocks may be as much as 10 ft (3 m) in maximum dimension. Block sizes decrease outward from the source.

Pumice occurs in all the western USA from the Rocky Mountains to the Pacific Ocean, and pumicite deposits are found in California, Texas, Oklahoma, Kansas, North and South Dakota, and Nebraska (nearly to the Missouri River). In 1997 at least 15 US mines operated, with Oregon being the leading pumice producer, followed by New Mexico, California, Idaho, and Arizona. Pumicite was produced in Kansas. Although the range in deposit size is large, a conservative estimate suggests that many occurrences have dimensions of 500 ft by 1000 ft by 10 ft (150 m by 300 m by 3 m). Bulk densities range from 850 lb yd^{-3} (500 kg m^{-3}) for pumice to 1750 lb yd^{-3} (1040 kg m^{-3}) for pumicite; amounts for the sample deposits lie between 80 000 and 160 000 short tons (7.3×10^4–14.5×10^4 t).

The total US pumice resource, based on the estimate of Higgs (1949), can only be approximated to between 150 and 300 million short tons (1.35×10^8–2.7×10^8 t). Great Plains pumicite deposits are numerous and vary greatly in size; Ham (1949) reported one in Oklahoma of about 1×10^7 yd^3 (7.6×10^6 m^3), and estimated about 2×10^7 yd^3 (15×10^6 m^3) in the known Oklahoma deposits. Carey *et al.* (1952) estimated about 20 million short tons (18×10^6 t) for Kansas. Data are lacking for Nebraska, but about the same tonnage as Kansas and Oklahoma is reasonable. A conservative estimate for all the Great Plains states is a total resource of about 100–150 million short tons (9×10^7–14×10^7 t).

1.5 Diatomite

Diatomite is composed of the siliceous skeletons of unicellular aquatic algae, both marine and freshwater, typically in the range 80–95% SiO_2. Although diatomite is used as a lightweight aggregate, the usage seldom exceeds a few tens of thousands of tons per year and is not a significant factor as a construction aggregate. Its value for other uses, such as filter aids, is far too great to justify its use as a low-cost aggregate. For information on origin, composition, etc., see Antonides (1999), Breese (1994), and Durham (1973).

1.6 Expanded Blast Furnace Slag

Expanded blast furnace slag is a small part of the supply of structural lightweight aggregates. Annual production since 1985 has been about 1 million short tons (9×10^5 t); in 1997 that was about 12% of the total lightweight aggregate construction-related usage. The amount used was only 0.6% of all blast furnace slag produced, and the supply seems ample for many years to come. The slag typically contains 32–42% SiO_2, 7–16% Al_2O_3, 32–45% CaO, 0.1–1.5% Fe_2O, 0.2–1.0% MnO, 0.5% MgO, and 1–2% S (R. S. Kalyoncu, personal communication, June 1999).

1.7 Fly and Bottom Ash

Fly ash and bottom ash, by-products of power plant coal combustion, are produced in large quantities (77.5 million short tons (7×10^7 t) in 1997) but only 22% is used for construction. From 1992 to 1997 total ash production increased steadily from about 60 to 77.5 million short tons (5.5×10^7–7×10^7 t) and the construction component followed. Average composition of fly and bottom ash is 40–55% SiO_2, 20–30% Al_2O_3, 6–10% Fe_2O_3, 8–20% CaO, ~2% MgO, 0–1% TiO_2, ~0.5% P_2O_5, 1–5% SO_2, and ~0.5% CO_2 (R. S. Kalyoncu, personal communication, June 1999). There probably is sufficient material for many years to come.

1.8 Exploitation

Pumice was used as construction material by the Romans about 2000 years ago. Cinders from coal-fired furnaces and boilers were used in cinder block from the late 1800s to World War II; conversion to gas and oil and the use of pulverized coal has virtually ended the cinder block industry. Fly ash and bottom ash, residues of pulverized-coal combustion, are used as pozzolans, aggregates, and starting materials for synthetic zeolites. In the USA expanded clay and shale aggregates were used as early as World War I in concrete ships, and a small market (partly owing to patent restrictions) persisted to the end of World War II. A rapid increase in US plants and production of expanded clay and shales began in about 1947 (patents on the rotary kiln (Hayde) process expired); production reached a high of more than 10.2 million short tons (9.3×10^6 t) in 1973 and remained steady at 4–5 million short tons (3.5×10^6–4.5×10^6 t) annually to 1997. Production and consumption of pumice and expanded blast furnace slag also increased very rapidly in the post-war construction boom of 1945–50. Since 1956 overall production of the structural lightweight aggregates has reflected construction trends fairly closely, although a long-term decline in the market share of expanded slag has occurred.

From 1964 to 1982 US scoria production was about 3 million short tons (2.7×10^6 t), then to 1997 about 2.5 million short tons (2.3×10^6 t). Pumice tonnage used as structural aggregate decreased from a 1965 high of 3 million short tons (2.7×10^6 t) to less than 55000 short tons (5×10^4 t) by 1997. Concurrently, usage in concrete block rose from about 30 000 short tons (2.7×10^4 t) in 1977 to about 375 000 short tons (3.4×10^5 t) in 1997.

Expanded slag production peaked near 3 million short tons (2.7×10^6 t) between 1955 and 1958, and has declined since then, hovering near 2 million short tons (1.8×10^6 t) annually since 1991. The decreases reflect a loss of markets to competitive products, particularly the expanded and sintered clays, shales, and slates. Sources of slag are chiefly iron blast furnaces in only a few areas of the USA. Suitable sources of clay, shale, and slate are widespread and new facilities compete with slag in market areas close to the furnaces, as well as in the areas where slag use is a marginal operation. Usage of fly and bottom ash has also contributed to the decline.

The lightweight aggregate industry is faced with the environmental problems of open-cast mining, the control of fine materials formed in the manufacturing process, the destructive consumption of water, and the degradation of water quality in other processes. These problems cause extra expense in production in a highly competitive market that has commonly slender profit margins. As low-unit-value, high-place-value commodities, raw material sources and production sites have to be close together and centrally located to the market area. Consumption of structural aggregates is mostly in urban areas except for highway construction; environmental impacts and conflicts are increased because they take place close to where many people live. One further complication is that in some areas the raw material must be extracted before the area is irretrievably lost to residential or industrial and office construction.

2. *Ultralightweight Aggregates*

The natural ultralightweight aggregates are perlite (product of silicic volcanism), vermiculite (a low-temperature alteration product of ultramafic igneous and metamorphic rocks), and, to a small degree, pumice and pumicite (described above). All require high-temperature heating to cause expansion and vesicularity (pumice is the exception). All develop high porosity, low permeability, and very thin cell walls. They make excellent thermal and acoustic insulators, but have low compressive strength in concrete and plaster (100 psi (7×10^5 Pa) to about 2200 psi (1.5×10^7 Pa)); more cement results in greater strength and higher weight. Ultralightweight aggregates are used in fireproof plasters, concrete curtain walls, floor fills, and roof decks, in refractory and insulating walls in bake ovens, and in decorative, nonload-bearing blocks and facings. In loose fill or bulk form they are used for insulation, but much is used for inert carriers and extenders in horticultural products and in manufacturing processes.

2.1 Volcanic Rocks

(*a*) *Perlite*. Perlite, obsidian, and pitchstone are volcanic glasses that differ in water content: obsidian, less than 2% H_2O; perlite, 2–5%; and pitchstone, more than 5%. The general range of composition (Bush 1973, p. 341) is 70–75% SiO_2, and about 10–15% Al_2O_3. The water content is critical to expansion, but obsidians with as little as 0.2% H_2O are reported to have expanded satisfactorily. Originally, perlite meant a glassy rock characterized by concentric cracks, but commercial usage includes any glassy rock (except pumicite) that can be expanded by heating to about 2000 °F (1100 °C).

Many of the traditional applications of perlite are in the construction field, but US usage decreased in 1997 to less than 5%. Most 1997 production went into formed products such as acoustic ceiling tiles, pipe insulation, roof insulation board, etc. (67%), filter aids (9%), and horticulture (10%). Diatomite is the principal material used in filtration, but perlite has made significant inroads into that market. In the construction field, the principal competition comes from exfoliated vermiculite: in 1997 some 30 000–40 000 short tons (2.7×10^4–3.6×10^4 t) perlite vs. 31 000 short tons (2.8×10^4 t) vermiculite. Breese and Barker (1994) give a comprehensive discussion of perlite processing, marketing, uses, and substitutes.

Perlite is found in three major associations: in glassy zones in welded ash-flow tuffs (Smith 1960), in silicic lava flows from volcanic domes, and in the wall zones of felsitic intrusive plugs and dikes. In a simple ash flow, the central zone of dense welding is a glass with virtually no porosity (composed of completely collapsed pumice and ash), bordered both above and below by zones of partial welding and no welding. Recurrent ash flows, above still hot flows, can form composite cooling units in which zones of dense welding may coalesce; intermediate partially welded and nonwelded zones are obliterated and a zone of glass borders and envelops a zone of tuff. If cooling is sufficiently rapid to prevent crystallization, the glass may be preserved for some millions of years. Preservation of glasses is increasingly rare from the Pliocene (1.6 million years ago) back to the Paleocene (about 60 million years ago) and is very rare in rocks older than that.

Glass in lava flows and volcanic domes represents the surface of rapidly cooled liquid lava which may be rolled under as the flow advances, in the process being brecciated, remelted, and annealed. As the flow stops, glass can be at both its top and its base, and if rapidly cooled, the zones may coalesce to make a virtually all glass flow.

Perlite very commonly surrounds microscopic to 1 in (2.5 cm) size cores of obsidian whose original water content is a few tenths of 1%. Around the cores is meteoric water at 2–5% due to hydration by groundwater, rain, and snow (Ross and Smith 1955; Hindman 1994and references therein). Increasing hydration outward and increasing devitrification inward are of importance to the assessment of perlite resources.

Perlite occurs worldwide. It is widely distributed throughout the western conterminous USA. In 1997 by far the most crude ore came from New Mexico; the remainder came from Arizona, California, Nevada, and Oregon. In earlier years crude ore came from Colorado, Idaho, and Texas, and deposits are known in Montana, Wyoming, and Washington.

Resources of perlite in the USA are ample and based on the very large volumes of available extrusive rocks, no shortage is foreseeable. There are very large deposits in New Mexico, and numerous sizable deposits in California, Arizona, Nevada, and smaller ones in Idaho, Utah, and Colorado. The total resource may be about 650 million short tons (6×10^8 t). However, hydration by meteoric water raises the problem of how far the glasses are hydrated inward from their margins. In the absence of data on the extent of hydration, the resource estimate must be considered a maximum.

(*b*) *Pumicite*. Laboratory investigations by the Oklahoma Geological Survey in 1944 revealed that pumicite was capable of considerable expansion upon heating in excess of 1100 °C. Additional investigations by the Kansas Geological Survey provided more data on temperatures and strengths and suggested industrial uses (filter aids and ultralightweight construction materials) (see also Sect. 1.4) Commercial usage of expanded pumicite was minor to 1999, but two operations are known from deposits in Kansas.

2.2 Vermiculite

Vermiculite is a mica, a hydrated, magnesium–iron–aluminum trioctahedral sheet silicate of varied composition. Its basic unit is two tetrahedral silicate sheets interlayered by an octahedral sheet that contains aluminum and iron (Grim 1962). Each tripartite layer, on its upper and lower surfaces, has regularly arranged "holes" that oppose corresponding sites on other layers. In biotite, potassium atoms occupy these sites; in vermiculite, only some sites are occupied, and these only by magnesium, but a double layer of water molecules also is present in these interlayer regions.

Vermiculite ore is invariably an alteration product of a variety of mafic minerals: biotite, phlogopite, rarely muscovite, diopside, tremolite, augite, hornblende, olivine, chlorite, and serpentine. All known deposits have formed in the zone of circulating groundwater above the water table. The alteration may include an intermediate biotite or chlorite stage, for some x-ray diffraction studies showed interlayering with biotite or chlorite, although no biotite or chlorite may remain in the host rock. Biotite is the common parent mineral; the similarity of crystal structures permits intimate interlayering on a molecular scale, in a series that ideally extends either regularly or irregularly from pure biotite through 1:1 biotite–vermiculite to pure vermiculite. The 10–50% vermiculite varieties are called hydrobiotite and are least likely to be useful commercially.

Virtually all vermiculite ores require beneficiation, by either dry comminution and air classification or by wet froth flotation methods, before exfoliation, almost always in vertical furnaces at temperatures commonly of 1700 °F (900 °C) or hotter for some ores. Gangue material is removed by air classification after exfoliation. Commercially, "vermiculite" includes any micaceous mineral that is capable of expansion. Pure vermiculite can be exfoliated up to 30 times its original thickness when heated to about 1600–2000 °F (870–1100 °C); impure varieties expand less. The exfoliation results from the flash conversion of interlayer water molecules into steam. High porosity develops as the layers are forced apart into an accordion-like structure. The resulting material is easily compressed along the line of expansion but has moderate strength in other directions.

Unexpanded vermiculite has very few uses, but exfoliated has many: in the construction industry (concrete as aggregate, in plaster and premix, and in blocks), in insulation (loose fill, acoustical tiles), in horticulture (aggregate, soil amendment, fertilizer carrier), and miscellaneous other uses. In 1990, the last year for which detailed US data are available, about 45% went to construction, 40% to agriculture, and 15% to insulation and other uses.

Vermiculite deposits occur worldwide and the largest are at large ultramafic intrusives, such as pyroxenite plutons, sometimes zoned, cut by syenite or alkalic granite and by carbonatitic rock and pegmatite (examples are deposits at Libby, Montana, and Phalaborwa, South Africa). Moderate size deposits are at small to large ultramafic intrusives, such as dunite and unzoned pyroxenite and peridotite cut by pegmatite and syenitic or granitic intrusive rock (examples are deposits in the Blue Ridge region of the southeastern USA). The smallest deposits are associated with ultramafic metamorphic rocks, commonly layered amphibolite schists that may have originally been igneous rocks, in places cut by or in contact with pyroxenite or peridotite, and cut by pegmatite (examples are deposits in the Enoree district, South Carolina, and most of the deposits in Colorado, Wyoming, and Texas). With the major exception of the Cretaceous deposit at Libby, almost all the vermiculite deposits in the USA appear to be in rocks of Precambrian age.

The geologic relations of all deposits suggest that pyroxenes, amphiboles, and olivine in the ultramafic rocks were altered by solutions and volatiles from intrusive syenite, carbonatite, and pegmatite (supplying mostly OH), and biotite, phlogopite, serpentine, and chlorite were formed. Supergene alteration by circulating groundwater later removed the alkalis, redistributed magnesium, and added H_2O as interlayer water molecules to form vermiculite.

The vermiculite deposits of Montana produced about three-quarters of the US crude ore to 1990 (cumulatively about 9 million short tons (8×10^6 t)). Since then the production is estimated to have come about equally from Virginia and South Carolina, although detailed data are unavailable. Production from all other states is very minor. The almost invariable association of vermiculite with Precambrian rocks allows Hawaii and virtually all the central part of the country from the Rocky Mountains east to the Blue Ridge to be dismissed from consideration, except possibly the Wichita Mountains and the Ozark Mountains.

Deposits of the Libby size, which would contain the largest resource, have geologic relations that are suggestive for additional discoveries. The association of ultramafic plutons with syenite, alkalic rock, and carbonatite is too common to be coincidence. The association of these rocks is known from several deposits in the USA (Rocky Boy, Montana; Gem Park-McClure Mountain and Powderhorn, Colorado), in Russia, in Brazil, and, in particular, in many parts of Africa. Several generalized clues to places where these rocks may be found together are (i) the rocks are usually in Precambrian terranes, (ii) they may be related to Precambrian structural trends, (iii) they may exhibit significant magnetic anomalies, and (iv) they may be characterized by ring-dike structures.

2.3 Exploitation

In the USA production trends of perlite and vermiculite were broadly similar before 1977, and very similar from about 1955 to 1992. To 1997 perlite usage increased whereas vermiculite was virtually unchanged. Perlite production began after World War II; output rose rapidly to 1951 and continued at a generally steady rate to more than 0.9 million short tons (0.8×10^6 t) in 1978. It has ranged from about 0.6 to 0.9 million short tons (0.5×10^6–0.8×10^6 t) since then, with about 0.72 million short tons (0.65×10^6 t) in 1997. The first sustained vermiculite production was in 1925, reached a high of 0.31 million short tons (0.28×10^6 t) in 1973, was about 0.3 million short tons (0.27×10^6 t) until 1988, and since has dropped to less than 0.2 million short tons (0.18×10^6 t). Since 1989 South Africa has been the world's largest source of crude vermiculite, in 1995 (the most recent year with detailed data) producing about 0.25 million short tons (0.22×10^6 t); the USA is second with about 0.2 million short tons (0.17×10^6 t) (US Geological Survey Minerals Yearbooks). Consumption of expanded perlite has varied more than that of exfoliated vermiculite, in part reflecting the closing of the major vermiculite deposit at Libby, Montana, and in part the movement of perlite into the filler, formed products, and filter-aids fields while vermiculite use decreased in concrete and plaster.

The USA produces all the perlite and pumicite that it consumes and has ample material to export, but, with its low unit value, only a few thousand tons of crude perlite is exported annually—all of it to Canada. No pumicite is exported. Almost all vermiculite used in the USA is produced domestically, but along the eastern seaboard, oceanic freight rates allow some vermiculite to be imported economically from South Africa.

Perlite, pumicite, and vermiculite share the common problems of open-cast mining. Large amounts of fines are produced in mining and processing, in the grinding circuits, and in the explosive expansion ("popping") of perlite and pumicite. These materials are glasses and the fines are pointed, sharp slivers and shards. Disposal and dust control are major problems, both in preventing and alleviating pollution and in protecting plant workers from respiratory injuries and diseases. Similar problems exist for vermiculite, which yields thin flakes to the fines rather than shards, but which may have asbestiform minerals in the gangue. Where wet classification methods are used, the disposal of vermiculite slimes presents other difficulties.

Bibliography

Antonides L E 1999 Diatomite. *Minerals Yearbook* 1997. US Geological Survey, Vol. 1

Austin C R, Nunes J L, Sullivan J D 1942 Basic factors involved in the bloating of clays. *Min. Technol.* **6** (4),

Boettcher A L 1966 Vermiculite, hydrobiotite, and biotite in the Rainy Creek igneous complex near Libby, MT. *Clay Miner.* **6** (4), 283–96

Boettcher A L 1967 The Rainy Creek alkaline-ultramafic igneous complex near Libby, Montana. *J. Geol.* **75** (5), 526–53

Breese R O Y 1994 Diatomite. In: Carr D D (ed.) *Industrial Minerals and Rocks*, 6th edn. Society of Mining Metallurgy and Exploration, Littleton, CO, pp. 397–412

Breese R O Y, Barker J M 1994 Perlite. In: Carr D D (ed.) *Industrial Minerals and Rocks*, 6th edn. Society of Mining Metallurgy and Exploration, Littleton, CO, pp. 735–49

Bush A L 1973 Lightweight aggregates. In: Brobst D A, Pratt W P (eds.) *United States Mineral Resources*, Prof. Paper 820. US Geological Survey, pp. 333–55

Durham D L 1973 Diatomite. In: Brobst D A, Pratt W P (eds.) *United States Mineral Resources*, Prof. Paper 820. US Geological Survey, pp. 191–5

Geitgey R P 1994 Pumice and volcanic cinder. In: Carr D D (ed.) *Industrial Minerals and Rocks*, 6th edn. Society of Mining Metallurgy and Exploration, Littleton, CO, pp. 803–13

Grim R E 1962 *Applied Clay Mineralogy*. McGraw-Hill, New York, pp. 20–37

Ham W E 1949 Volcanic ash deposits in Oklahoma. In: Burwell A L (ed.) *Cellular Products from Oklahoma Volcanic Ash*, Circ. 27. Oklahoma Geological Survey, pp. 42–81

Higgs D V 1949 Quantitative areal geology of the United States. *Am. J. Sci.* **247** (8), 575–83

Hindman J R 1994 Vermiculite. In: Carr D D (ed.) *Industrial Minerals and Rocks*, (6th edn). Society of Mining Metallurgy and Exploration, Littleton, CO, pp. 1103–11

Hunter C E 1950 Vermiculite of the southeastern states. In: Snyder F G (ed.) *Symp. Mineral Resources of the Southeastern United States*. Tennessee University Press, Knoxville, TN, pp. 120–7

Mason B H 1994 Lightweight aggregates. In: Carr D D (ed.) *Industrial Minerals and Rocks*, 6th edn. Society of Mining Metallurgy and Exploration, Littleton, CO, pp. 343–50

Riley C M 1951 Relation of chemical properties to the bloating of clays. *Am. Ceram. Soc. J* **34** (4), 121–8

Ross C S, Smith R L 1955 Water and other volatiles in volcanic glasses. *Am. Mineral.* **40** (11/12), 1071–89

Schilling J.H. 1960 Mineral resources of Taos County, New Mexico. New Mexico Bureau of Mines and Mineral Resources, Bull. 71

Smith R L 1960 Zones and zonal variations in welded ash flows US Geological Survey Prof. Paper 354-F, pp. 149–51

Sullivan J D Austin C R Rogers E J 1942 Expanded clay products. American Institute of Mining and Metallurgy Engineers Tech. Pub. 1485

Sweeney J W Hamlin H P 1965 Lightweight aggregates—expansion properties of selected Illinois shales and clays. US Bureau of Mines Rep. Inv. 6614

A. L. Bush
Lakewood, Colorado, USA

Continuous Parallel Fiber Composites: Deformation and Strength

This article is concerned with the deformation and failure processes that occur in continuous unidirectional (UD) composites in tension, compression, and shear. Methods for calculating thermoelastic constants and strengths in both the longitudinal and transverse fiber directions are presented. The models for predicting longitudinal tensile strength allow for both weak and strong interface bonding between fibers and the matrix. The higher level of coverage given to thermoelastic properties and longitudinal tensile strength reflects the level of understanding and confidence in the predictive analysis for these properties.

The main failure modes associated with longitudinal tensile loading are fiber fracture and matrix cracking normal to the fiber direction. Fiber fracture is often accompanied by debonding of the fiber–matrix interface in the neighborhood of the fiber fractures. Three types of behavior need to be considered when modeling UD composites. At relatively low applied loads, there is no microstructural damage or matrix yielding/plasticity. The composite deforms as a linear material (i.e., stresses and strains are linearly related). All interfaces between fibers and matrix are perfectly bonded. Many models predicting the thermoelastic properties of UD composites have been developed for this type of behavior (see *Composites, Physical Properties of**). For larger applied loads, the composite experiences the second type of behavior where stable nonlinear deformation occurs, arising from matrix yield and/or from microstructural damage formation. Following the progressive stable growth of plasticity and/or microstructural damage, the third type of behavior ensues, where instability and catastrophic failure of the composite occur.

The results to be presented can be applied to any fiber-reinforced composite whose constituents can be regarded as deforming in a linear manner, such as thermosetting polymer matrix composites (PMCs) and fiber-reinforced ceramic matrix composites. For monotonically loaded, fiber-reinforced thermoplastic and metal composites, for which the matrix may undergo plastic deformation or creep, extensions of these analyses also exist; however, for clarity in presentation, we will place emphasis on the case of an elastic matrix. In general, it is assumed for UD composites that the fiber and matrix are made of different materials (possibly anisotropic, having directional properties), and that in the undamaged state they are perfectly bonded everywhere along the length of the fiber–matrix interfaces. The cross-section of the fibers is assumed to be uniform and circular. The fibers are assumed to be uniformly distributed, having a hexagonal packing structure leading to transverse isotropic composite properties.

1. Prediction of Thermoelastic Constants for Undamaged Composites

This section presents suitable methods for calculating the thermoelastic constants of undamaged UD composites, some of which are almost impossible to measure. Concentric cylinder models of fiber and matrix in conjunction with variational techniques have been used extensively to predict the properties of undamaged UD composites (Hashin 1983). UD composites are realistically modeled as transverse isotropic solids (i.e., axial properties differ from properties in directions normal to the fiber direction). UD composites are characterized by seven independent thermoelastic constants: (i) axial Young's modulus (E_{A}); (ii) axial shear modulus (μ_{A}); (iii) transverse Young's modulus (E_{T}); (iv) axial Poisson's ratio (ν_{A}); (v) transverse Poisson's ratio (ν_{T}); (vi) axial thermal expansion coefficient (α_{A}); and (vii) transverse thermal expansion coefficient (α_{T}). These parameters appear in stress–strain relations of the form:

$$\varepsilon_{rr} = \frac{1}{E_{\mathrm{T}}}\sigma_{rr} - \frac{\nu_{\mathrm{T}}}{E_{\mathrm{T}}}\sigma_{\theta\theta} - \frac{\nu_A}{E_{\mathrm{A}}}\sigma_{zz} + \alpha_{\mathrm{T}}\Delta T \qquad (1)$$

$$\varepsilon_{\theta\theta} = -\frac{\nu_{\mathrm{T}}}{E_{\mathrm{T}}}\sigma_{rr} + \frac{1}{E_{\mathrm{T}}}\sigma_{\theta\theta} - \frac{\nu_{\mathrm{A}}}{E_{\mathrm{A}}}\sigma_{zz} + \alpha_{\mathrm{T}}\Delta T \qquad (2)$$

$$\varepsilon_{zz} = -\frac{\nu_{\mathrm{A}}}{E_{\mathrm{A}}}\sigma_{rr} - \frac{\nu_{\mathrm{A}}}{E_{\mathrm{A}}}\sigma_{\theta\theta} + \frac{1}{E_{\mathrm{A}}}\sigma_{zz} + \alpha_{\mathrm{A}}\Delta T \qquad (3)$$

$$\varepsilon_{rz} = \frac{\sigma_{rz}}{2\mu_{\mathrm{A}}}, \varepsilon_{\theta z} = \frac{\sigma_{\theta z}}{2\mu_{\mathrm{A}}}, \varepsilon_{r\theta} = \frac{\sigma_{r\theta}}{2\mu_{\mathrm{T}}} \qquad (4)$$

The stress σ and strain ε components refer to cylindrical polar coordinates (r, θ, z), with the z-axis oriented parallel to the fibers.

1.1 Method for Estimating Properties

For an undamaged UD composite having fiber and matrix volume fractions V_{f} and V_{m} $(=1-V_{\mathrm{f}})$, respectively, the axial Young's modulus, axial Poisson's ratio, and transverse bulk modulus are given by:

$$E_{\mathrm{A}} = V_{\mathrm{f}}E_{\mathrm{A}}^{\mathrm{f}} + V_{\mathrm{m}}E_{\mathrm{A}}^{\mathrm{m}} + 2\lambda(\nu_{\mathrm{A}}^{\mathrm{m}} - \nu_{\mathrm{A}}^{\mathrm{f}})^2 V_{\mathrm{f}}V_{\mathrm{m}} \qquad (5)$$

$$\nu_{\mathrm{A}} = V_{\mathrm{f}}\nu_{\mathrm{A}}^{\mathrm{f}} + V_{\mathrm{m}}\nu_{\mathrm{A}}^{\mathrm{m}} - \frac{\lambda}{2}(\nu_{\mathrm{A}}^{\mathrm{f}} - \nu_{\mathrm{A}}^{\mathrm{m}})\left(\frac{1}{k_{\mathrm{T}}^{\mathrm{f}}} - \frac{1}{k_{\mathrm{T}}^{\mathrm{m}}}\right)V_{\mathrm{f}}V_{\mathrm{m}} \qquad (6)$$

$$\begin{aligned}\frac{1}{k_{\mathrm{T}}} \equiv \frac{2(1-\nu_{\mathrm{T}})}{E_{\mathrm{T}}} - \frac{4\nu_{\mathrm{A}}^2}{E_A} &= \frac{V_{\mathrm{f}}}{k_{\mathrm{T}}^{\mathrm{f}}} + \frac{V_{\mathrm{m}}}{k_{\mathrm{T}}^{\mathrm{m}}} \\ &- \frac{\lambda}{2}\left(\frac{1}{k_{\mathrm{T}}^{\mathrm{f}}} - \frac{1}{k_{\mathrm{T}}^{\mathrm{m}}}\right)^2 V_{\mathrm{f}}V_{\mathrm{m}} \qquad (7)\end{aligned}$$

where

$$\frac{1}{\lambda} = \frac{1}{2}\left[\frac{1}{\mu_{\mathrm{T}}^{\mathrm{m}}} + \frac{V_{\mathrm{f}}}{k_{\mathrm{T}}^{\mathrm{m}}} + \frac{V_{\mathrm{m}}}{k_{\mathrm{T}}^{\mathrm{f}}}\right] \qquad (8)$$

E, ν, k, μ, and α denote Young's modulus, Poisson's ratio, bulk modulus, shear modulus, and thermal coefficient of expansion, respectively. Suffixes f and m denote fiber and matrix, respectively. As the values of E_{A} and ν_{A} are known, the value of $(1-\nu_{\mathrm{T}})/E_{\mathrm{T}}$ can be calculated using Eqn. (7). Assuming that the composite is transversely isotropic, then the following relationship holds:

$$E_{\mathrm{T}} = 2\mu_{\mathrm{T}}(1+\nu_{\mathrm{T}}), \quad \text{but} \quad E_{\mathrm{A}} \neq 2\mu_{\mathrm{A}}(1+\nu_{\mathrm{A}})$$

Clearly both E_{T} and ν_{T} can be calculated once the value of μ_{T} has been determined. Christensen and Lo

(1979) developed a concentric cylinders model (including a surrounding effective composite region) that can be used to estimate transverse shear modulus, which lies between the bounds given by Hashin (1983), and is likely therefore to predict more accurate values. The model can be used to calculate the transverse shear modulus μ_T, thus enabling values of E_T and ν_T to be determined. The following expression given by Hashin (beware a misprint in his paper), leads to either an upper or lower bound value for μ_T, depending upon the values of μ_f, μ_m, k_f, and k_m:

$$\mu_T = V_f \mu_T^f + V_m \mu_T^m - \lambda_1 (\mu_T^f - \mu_T^m)^2 V_f V_m$$

where

$$\frac{1}{\lambda_1} = V_m \mu_T^f + V_f \mu_T^m + \frac{k_T^m \mu_T^m}{k_T^m + 2\mu_T^m} \tag{9}$$

Hashin's (1983) expression for the axial shear modulus μ_A may be written as:

$$\mu_A = V_f \mu_A^f + V_m \mu_A^m - \lambda_2 (\mu_A^m - \mu_A^f)^2 V_f V_m$$

where

$$\frac{1}{\lambda_2} = \mu_A^m (1 + V_f) + \mu_A^f V_m \tag{10}$$

The axial and transverse thermal expansion coefficients predicted by the model may be written as:

$$\begin{aligned} E_A \alpha_A = {} & V_f E_A^f \alpha_A^f + V_m E_A^m \alpha_A^m + 2\lambda (\nu_A^m - \nu_A^f) V_f V_m \\ & \times [\alpha_T^m + \nu_A^m \alpha_A^m - \alpha_T^f - \nu_A^f \alpha_A^f] \end{aligned} \tag{11}$$

$$\begin{aligned} \alpha_T = {} & -\nu_A \alpha_A + (\alpha_T^f + \nu_A^f \alpha_A^f) V_f + (\alpha_T^m + \nu_A^m \alpha_A^m) V_m \\ & + \frac{\lambda}{2} \left(\frac{1}{k_T^f} - \frac{1}{k_T^m} \right) V_f V_m \\ & \times [(\alpha_T^m + \nu_A^m \alpha_A^m) - (\alpha_T^f + \nu_A^f \alpha_A^f)] \end{aligned} \tag{12}$$

The relations in Eqns. (10) to (12) are equivalent to those given by Hashin (1983) and Rosen and Hashin (1970). For composites where fiber coatings or reaction products (e.g., fiber-reinforced titanium) are encountered, thermoelastic constants can be predicted using a multiple concentric cylinder model (McCartney 1992a). This model can deal with any number of transverse isotropic concentric layers.

2. *Deformation of Undamaged Composites in the Presence of Matrix Yield*

Thermoplastic and metal matrices deform by (permanent) plastic deformation once a certain stress is exceeded: when unloaded, composites with such matrices can show a permanent deformation even in the absence of damage. Hill (1964) has explored extensions of elastic modulus bound predictions for the deformation of elastoplastic fiber composites, and has in particular shown that for uniaxial composite deformation parallel to the fibers, the slope of the stress–strain curve is close to the equistrain rule of mixtures (as is the case for elastic deformation, Eqn. (5) above). For this case, then, prediction of the composite deformation laws is relatively straightforward, and can be extended to include other matrix rheologies (such as creep), and also to take into account the influence of residual stresses caused by differential matrix–fiber cooling during cool-down from composite processing temperatures.

For all other modes of composite straining in the presence of matrix plasticity, prediction of the composite flow stress becomes complicated. This is in large part because the matrix strain distribution is highly inhomogeneous in such materials and can essentially only be predicted using numerical analysis methods. Finite-element micromechanical models for the deformation of fiber-reinforced elastoplastic matrices are described by Dvorak (1991), Suresh and Brockenbrough (1993), and Clyne and Withers (1993), together with approximate analyses based on extensions of elasticity models. Overall, numerical models show good agreement with experiment and account relatively well for phenomena such as the strong Bauschinger effect that is observed with such composites, or the influence of residual thermal stresses. The main outstanding difficulty in modeling the deformation of elastoplastic fiber composites is that, contrary to elasticity, the intrinsic rheology of the matrix may be affected by the fibers: in metal matrix composites, fibers a few micrometers in diameter are on the same length scale as those microstructural elements that govern plastic deformation in metals. Hence, it is quite possible that the matrix flow stress will be affected by the fibers, and will differ from its value in the unreinforced state. This is indeed observed: the flow stress of fiber-reinforced metals has been found to depend on the reinforcement diameter (Pedersen 1990), something that is not predicted by (scale-independent) elastoplastic mechanical models.

3. *Failure Processes during Axial Tensile Loading*

The behavior of UD composites when loaded in the fiber direction can be dominated by the properties of the fibers, especially for relatively high volume fractions of high-modulus fibers in a comparatively low-modulus matrix. This is particularly the case for carbon fiber-reinforced epoxy composites, where the high modulus of the fibers and low modulus of the matrix results in the fibers supporting most of the applied load. For metal and ceramic matrix composites where the matrix has much higher moduli than for polymers, the fibers do not carry as much load;

however, the difference in stress between fiber and matrix often remains high (and ncreases if the matrix yields by plastic deformation). The strengths of UD composites, as with their moduli, are very anisotropic. The strength transverse to the fiber direction is strongly affected by both matrix and fiber–matrix interface properties, and also by potential stress concentration sites at free surfaces perpendicular to the fibers. Its reliable prediction is therefore very difficult.

3.1 Statistics of Fiber Strength

The majority of the literature regards materials strength as a random variable, without inquiring too deeply into the causes of material failure, and most statistical treatments of failure adhere strictly to the weakest link failure hypothesis. A chain breaks in tension when one or more of the links fail. The failure event is thus controlled by the weakest link (or links) in the chain. The failure of a fiber is perhaps the simplest application for weakest link statistics (see *Fiber Strength**). The fiber length is considered to consist of N elements of equal length that are small enough for just one defect to be present. The strength of such elements is a random variable. The assembly of elements (i.e., the whole fiber) breaks when the load carried by the fiber is large enough for the weakest element (containing the largest defect) to fail. Just one element needs to fail to cause the failure of the fiber. Suppose that $F(\sigma)$ is the cumulative strength distribution for one element of the system, i.e., $F(\sigma)$ is the probability that the strength of the element is less than σ, so that $1-F(\sigma)$ is the probability that the strength is greater than σ.

For an assembly of N elements arranged as a chain (as would be the case for a fiber), each with statistically independent strength characteristics and each subjected to a uniform stress σ, the probability that the assembly survives is $[1-F(\sigma)]^N$, since each element must have strength greater than σ. The probability that the application of the stress σ causes the failure of at least one element is then given by:

$$P_{\mathrm{f}}(\sigma) = 1 - [1 - F(\sigma)]^N \tag{13}$$

which is a mathematical statement of the weakest link failure hypothesis. When the number of fiber elements N becomes very large, as their length decreases, it is possible to apply the asymptotic results of extreme value theory (Gumbel 1954). In the asymptotic limit $N \to \infty$ there are just three statistical distributions that arise from Eqn. (13). Two of these involve a statistical distribution for strength σ defined over the interval $-\infty < \sigma < \infty$, which is not suitable for applying to strength measurements because the strength of fibers can be sensibly defined only for positive values. The third possible asymptotic distribution, the Weibull distribution (1951), is the only asymptotic distribution that is defined on the interval $0 < \sigma < \infty$. Thus, there are sound statistical reasons, based on extreme value theory, for using the Weibull distribution when correlating experimentally measured values of fiber strength. The cumulative form of the Weibull distribution is:

$$P_{\mathrm{F}}(\sigma) = 1 - \exp[-\delta(\sigma/\sigma_{\mathrm{o}})^m] \tag{14}$$

where $P_{\mathrm{F}}(\sigma)$ is the probability of failure in the stress range $0 \to \sigma$ and where δ is the length of each fiber element. The parameter m is the Weibull modulus and σ_{o} is a scaling parameter.

3.2 Strength of Fiber Bundles and UD Composites with Weakly Bonded Interfaces

For PMCs where interface bonding is poor, to the extent that it can be neglected, the UD composite, when loaded in the fiber direction, behaves in a similar way to a loose bundle of fibers where *equal load sharing* (or global load sharing) can be assumed following fiber fractures. For such composites, a fiber failure means that the failed fiber is unable to support load at any point along its length and the load it carried is shared equally amongst all surviving fibers. The classical work of Daniels (1945) showed that, as the number of fibers increased, the statistical distribution for bundle strengths is asymptotic to a normal distribution. The analysis provides an accurate value for the mean strength of the bundle and it indicates that the standard deviation tends to zero as the number of fibers in the bundle tends to infinity.

The important result is that if the number of fibers in a bundle is large enough, then the strength of the bundle has a determinate value that is easily derived from the strength distribution for individual fibers, namely $\sigma_{\mathrm{o}} m^{-1/m}$. Coleman (1958) applied the general theory of Daniels to the case when the fiber strengths are governed by the Weibull distribution. The early analyses were not able to indicate just how large a bundle must be before its strength distribution can be represented by the normal distribution. McCartney and Smith (1983) carried out extensive, computer-based calculations showing that the strength distribution could be accurately approximated by a normal distribution whenever the number of fibers in a bundle exceeded 200. The strength properties of the fibers have an effect on this number. It is worth noting that for some ceramic composites in the last stages of damage development before final catastrophic failure, the interfaces are all debonded and the cracked matrix is behaving as though it is detached from the fibers. The composite strength derives solely from the fibers, with the result that catastrophic failure of such composites can be modeled using loose-bundle theory.

3.3 Strength of UD Composites with Strong Interface Bonding

Most PMCs have reasonably strong interfacial bond strength between fibers and matrix, leading to the situation that a failure in one fiber does *not* mean that the failed fiber cannot support load in regions along its length which are remote from the point of fiber fracture. Strong interfacial bonding also leads to the action of shear tractions at the interface in the stress transfer zone that is located in the neighborhood of the fiber break. Fiber fracture then results in local transfer of load from the failed fiber to the surrounding matrix and nearby intact fibers. This is known as *local load sharing* and occurs only within the *stress transfer length* (usually a few fiber diameters) associated with a fiber fracture or matrix crack.

Statistical models have been developed that allow load carried by a broken fiber to be shared equally amongst only the neighboring fibers. These models are applied only to very small lengths of composite material, of the order of several fiber diameters, representing the region where fiber–matrix stress transfer occurs. As macroscopic lengths of composite need to be considered, the models have been modified so that the composite is represented by series of bundles of fibers, where each bundle has the length of the stress transfer zone and is subject to local load sharing. These models are often referred to as series/parallel models that can be analyzed for both equal and local load sharing when fibers break (Smith and Phoenix 1981, Smith 1982, Phoenix and Smith 1983, Smith *et al.* 1983).

When the interface is strong, a fiber fracture occurs with very little fiber–matrix debonding. Stress transfer from the broken fiber to the neighboring matrix and surrounding fibers is localized and involves significant stress concentrations in the matrix and some stress concentration in neighboring unbroken fibers. This type of interface behavior leads to the growth of a dominant crack from the first fiber to fail. Fracture surfaces of the composite are almost smooth in nature and, contrary to what is observed with weak interfaces, there is little fiber pullout. For strongly bonded composites, stress transfer at broken fibers or matrix cracks can be modeled, assuming that the interface is perfectly bonded for all stress states.

For this case, *shear lag* modeling (a rather approximate technique for predicting stress transfer between fiber and matrix due to the action of shear tractions on the interface) was first proposed by Cox (1952), and further developed by Aveston and Cooper (1973) and Nayfeh (1977). A variational method (based on minimizing an energy function for an assumed approximate representation of the stress field) predicting the stress field was provided by Nairn (1992) and complete solution methods, giving both stress and displacement fields, have been developed by McCartney (1989), 1999a, Extensions of global and local load sharing analyses to the case of an elastoplastic or creeping matrix can be found in Beyerlein and Phoenix (1996), Foster *et al.* (1998), and Iyengar and Curtin (1997).

3.4 Characterising Interfacial Debonding

The measurement of fiber–matrix interface properties is usually carried out using special tests. Examples include the micro-drop test (suitable for polymer composites), fiber fragmentation test (suitable for matrices with high failure strain), fiber pull-out test (suitable for polymer composites), and fiber push-through test (suitable for metal and ceramic composites reinforced with large-diameter fibers). Each of these tests is designed to lead to fiber–matrix debonding, and the interpretation of the results of such experiments requires the use of a suitable model that enables interface properties to be related to the applied load and corresponding displacement. Interface "strength" is a very difficult concept to define unambiguously.

If one imagines that a fiber–matrix interface is subject to a uniform tensile radial stress, then one can accept that at a critical value of this radial stress the interface will fail, although failure may initiate from an interfacial defect. This type of interface failure does not usually occur in practice except from the effects of thermal residual stresses, where thermal expansion mismatch between fiber and matrix can in some cases lead to a sufficiently large uniform tensile stress at the interface. For the interface tests mentioned above, the loading of the interface is much more complex, involving combined shear and normal tractions and the influence of friction for regions of debonded interfaces.

Two distinct approaches have been made to modeling the complex stress states occurring during interface failure. The first approach assumes that the interface fails at a critical value of the interfacial shear stress τ, and that in the debonded region the interfacial shear stress has the uniform value τ (refer to the shear lag model of Aveston *et al.* 1971). The second approach assumes that stress transfer from fiber to matrix in the debonded region is governed by the Coulomb friction law, where the coefficient of friction is the interface parameter rather than τ (see McCartney 1989, Hutchinson and Jensen 1990). For this case the shear stress is nonuniform in the debonded region.

The constant shear stress model and the Coulomb friction law do not involve the interfacial energy associated with the progressive growth of an interfacial debond during loading. This energy can be included in the model only if the singularity at the debond tip is modeled correctly. This is very difficult to achieve using analytical modeling methods (see Tandon and Pagano 1996, 1998, Pochiraju *et al.* 1999). Analytical modeling is possible, however, if one ignores frictional effects resulting from debonding and considers only the energetics of propagation for a debond that is long enough for there to be negligible interaction between the debond tip and the crack in the fiber or matrix. For a frictionally bonded composite, where there is mechanical contact and negligible bonding between fiber and matrix, the use of the Coulomb friction model of interfacial behavior is justified.

3.5 Matrix Cracking and Bridged Cracks

Aveston *et al.* (1971) applied a shear lag model of fiber–matrix stress transfer to the prediction of matrix cracking in brittle matrix composites using an energy balance approach. The fiber and matrix strengths are such that all fibers remain intact when the matrix is traversed entirely by a crack. The well-known classical result for the matrix cracking stress $\varepsilon_{\mathrm{mc}}$ is:

$$\varepsilon_{\mathrm{mc}} = \left(\frac{12\gamma_{\mathrm{m}}\tau}{R}\,\frac{V_{\mathrm{f}}^2 E_{\mathrm{f}}}{V_{\mathrm{m}} E_{\mathrm{m}}^2 E_{\mathrm{A}}}\right)^{1/3} \tag{15}$$

where $2\gamma_{\mathrm{m}}$ is the fracture energy of the matrix, R is the fiber radius, and where E_{A} is the axial modulus of the composite in an undamaged state. Marshall *et al.* (1985), Budiansky *et al.* (1986), and McCartney (1987) have adopted a fracture mechanics approach to such predictions by considering the loading conditions that will cause a pre-existing defect in the matrix (bridged by fibers) to propagate unstably.

Key aspects are: (i) modeling the crack bridging relationship between the effective traction applied to the crack faces by intact fibers and the crack opening displacement; and (ii) the development of the energy release rate for bridged cracks in terms of the crack bridging relationship. Analyses of bridged cracks have shown that as the crack length increases the energy release rate becomes independent of the crack length, provided that fibers remain intact during loading. McCartney (1987), 1992b, 1992c), showed that Eqn. (15) was consistent with the fracture mechanics approach to the prediction of matrix cracking, provided that bridged matrix cracks are sufficiently long. Marshall and Cox (1987), Cox and Marshall (1991), and Cox (1991) have analyzed in detail the behavior of bridged cracks for the case when fiber fracture occurs during propagation. McMeeking and Evans (1990) and McCartney (1996) have modeled the growth of matrix cracks during fatigue loading. Cox and Rose (1994) have considered the modeling of crack bridging for the case of time and cyclic loading.

3.6 Monte Carlo Models of Tensile Strength

The modeling of the strength of a well-bonded composite using statistical principles has required simplifications that enable the analysis to be tractable. Statistical approaches to static strength prediction using equal or local load-sharing concepts with series/parallel structures are very rudimentary when dealing with stress transfer from fiber to matrix caused by fiber fracture. Imposing more realistic load-sharing methods leads to intractable statistical analysis. To overcome this problem and also to test statistical models that have been developed, Monte Carlo concepts have been used for static strength prediction. The Monte Carlo approach does not attempt to formulate and solve the problem of failure prediction using probability distribution functions and analytical techniques. The approach is to set up a numerical simulation of the failure of a UD composite, where the composite is regarded as an assembly of thin layers of equal thickness δ, having the same order as the stress transfer length.

Each layer is regarded as an array of representative volume elements of the composite, where each element comprises just one fiber element together with the matrix with which the fiber element is associated. This representative volume element can be thought of as a fiber–matrix cell. The failure strain of any fiber element in a cell is randomly generated from some statistical distribution function that characterizes their strength. Having attributed a failure strain to each fiber element, the model of the composite can be used to determine the failure strain of the composite by identifying the locations of fiber–matrix cells in the system where fibers fail and redistributing the load carried by a failing cell to surrounding fiber–matrix cells. For a given set of failure strains for the fiber elements in the system, the model predicts a single failure strain for the composite. The mean strength and statistical variability of composite strength can be determined only by carrying out many repeated simulations of composite failure.

Curtis (1986) developed a model that was designed to simulate the behavior of the composite by representing the composite as a set of stacked layers of rectangular arrays of fibers. For the Curtis model, the effects of the matrix are neglected and failure is assumed to occur when all of the fiber elements have failed in just one of the layers. A consequence of this assumption is that the fracture surface will be very flat and normal to the direction of the applied load. This would only be appropriate for composites where the interface strength is high. Curtis assumes that the failure strains of fiber elements can be modeled by the normal distribution. More recently, Curtin and Takeda (1998a) have developed a Monte Carlo model for the prediction of the tensile strength of UD fiber-reinforced composites in which stress transfer associated with fiber breaks is accounted for by a shear-lag model that imposes a uniform shear stress in the debonded region. Curtin and Takeda (1998b) have applied their model to investigate the strength behavior of polymer composites. Curtin *et al.* (1998)

developed the model further to allow for the effects of matrix cracking in brittle matrix composites.

The effective stress–strain behavior of the failed cell is an important input to any Monte Carlo simulation technique. This can be modeled through the use of a micromechanical model (McCartney 1999a) that accounts for localized load transfer between fiber and matrix associated with fiber failure. The model also takes account of the effects of debonding at the fiber matrix interface. Debonding will have a significant effect on the local stress transfer between fiber and matrix, leading to a significant effect on the effective axial stiffness of the fiber–matrix cell. The change in cell stiffness arising from a fiber fracture and associated fiber–matrix debonding leads to stress transfer between neighboring cells where the stress carried by the failed cell is reduced and that carried by the neighboring cells is increased, according to a specified load-sharing rule.

3.7 Longitudinal Compressive Strength

Failure of continuous, aligned, fiber-reinforced laminae under uniaxial compressive load may occur as a result of three competing mechanisms: (i) fiber microbuckling (or fiber kinking); (ii) Euler macrobuckling; and (iii) interlaminar shear failure. Both Euler macrobuckling and interlaminar shear failure can be avoided by careful consideration of the specimen dimensions (Soutis 1991). It is thought that microbuckling in composite laminates initiates as a result of the elastic bending of fibers, loaded by the matrix in shear (Weaver and Williams 1975). At a critical strain, the fibers fail in compression on their concave side at points of maximum fiber curvature. The microbuckle then spreads across the specimen width by a process of successive buckling and fracture along the fibers and by the transfer of load to adjacent fibers which have already buckled under the axial strain (see Fig. 1).

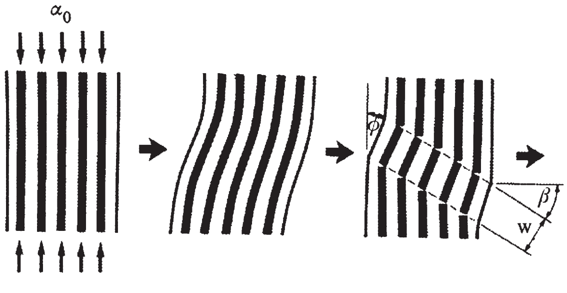


Figure 1
Fiber microbuckling under compressive loading (β = boundary orientation angle, ϕ = inclination angle (5–10°), and w = kink band width (10–15 fiber diameters)).

The process can be described in three stages: (i) elastic kinking; (ii) plastic kinking; and (iii) final collapse as a result of the fiber failure at the kink boundaries. The elastic kinking is initiated at the free edges of the laminate or at a defect, such as a void, fiber curvature or interfacial debond. The plastic kinking stage consists of plastic shear deformation of the matrix material under compression load. There is no definitive theoretical model for predicting longitudinal compressive strength, $S_{11}(\mathrm{C})$, of continuous aligned fiber-reinforced laminates. However, combining the modeling of Weaver and Williams (1975) and Budiansky (1983) it is possible to obtain a reasonable estimate of the longitudinal compressive strength. The critical composite stress can be calculated using the Euler buckling equation (Soutis 1991):

$$S_{11}(\mathrm{C}) = \left(\frac{\pi d_\mathrm{f}}{4w}\right)^2 E_{11} \tag{16}$$

where d_f is the fiber diameter, w is the kink band width (typically 10–15 fiber diameters), and E_{11} is the longitudinal elastic modulus of the composite (see Fig. 1). The fiber diameter and kink band width are related as follows (Soutis 1991):

$$\frac{w}{d_\mathrm{f}} = \frac{\pi}{4}\left(\frac{2\tau_\mathrm{m}^\mathrm{u}}{E_{11}}\right)^{-1/3} \tag{17}$$

Combining Eqns. (16 and 17) gives:

$$S_{11}(\mathrm{C}) = (2\tau_\mathrm{m}^\mathrm{u})^{2/3}(E_{11})^{1/3} \tag{18}$$

Recent modeling attempts to include fiber waviness, which is experimentally determined (Berbinau *et al.* 1999). These models are still under development.

3.8 Transverse Compressive Strength

An aligned, continuous, fiber-reinforced laminate under transverse compressive load usually fails by shear failure of the matrix (Fig. 2). This mode of failure may be accompanied by fiber–matrix debonding and crushing. Transverse compressive strength $S_{22}(\mathrm{C})$ for these materials is generally higher than the stress at which the unreinforced matrix material yields under compressive loading. This is due to the difficulty of fibers sliding across each other along the shear surfaces. Based on Mohr's strength theory, Skudra (1985) suggests that $S_{22}(\mathrm{C})$ can be determined from the following relation:

$$S_{22} = \frac{2S_{22}(\mathrm{T})\sin 2\phi}{(1+\nu_{23})(1-\cos 2\phi)} \tag{19}$$

where $S_{22}(\mathrm{T})$ is the transverse tensile strength of the lamina, ϕ is inclination angle of the fracture path (experimentally determined) with respect to the direction of loading (see Fig. 2). The value of ϕ is typically

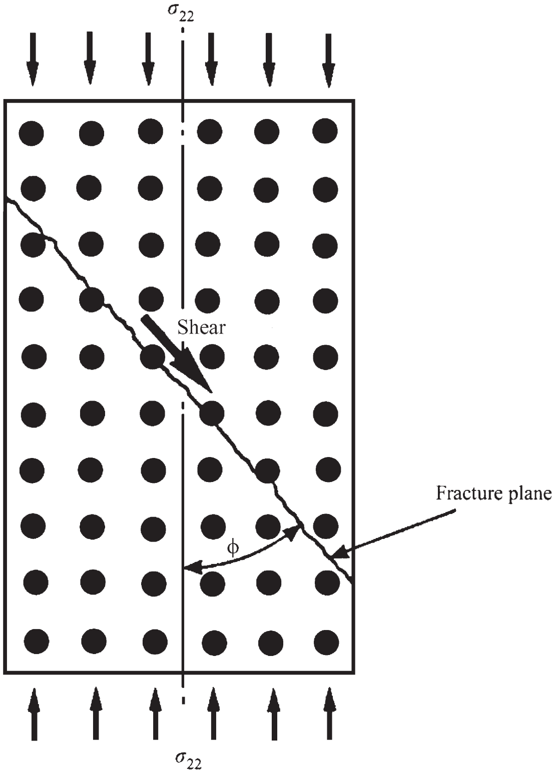


Figure 2
Transverse compressive fracture.

30° for a wide range of fiber-reinforced plastics (e.g., glass and carbon fiber-reinforced epoxy).

3.9 In-plane Longitudinal Shear Strength

Micromechanics modeling of in-plane shear strength S_{12} has not developed to the stage where a reliable prediction based on constituent material properties is possible. The simplest method of estimating shear strength is to assume that a "rule of mixtures" applies:

$$S_{12} = \tau_{\mathrm{i}}^{\mathrm{u}} A_{\mathrm{i}} + \tau_{\mathrm{m}}^{\mathrm{u}}(1 - A_{\mathrm{i}}) \tag{20}$$

where A_{i} is the fraction of the fracture area occupied by the interface, and $\tau_{\mathrm{i}}^{\mathrm{u}}$ and $\tau_{\mathrm{m}}^{\mathrm{u}}$ are the interfacial and matrix shear strengths, respectively. This approach assumes that the fracture path is planar and bisects a regular square array. Fracture usually propagates preferentially through the matrix in a nonplanar manner without disturbing the fibers, not conveniently along the fiber–matrix interface as assumed in Eqn. (20). Shu and Rosen (1967) performed a more rigorous analysis of the unidirectional laminae and obtained "lower" and "upper" bounds to S_{12}.

It was assumed that the fibers were elastic-brittle and rigid and the matrix was elastic-perfectly plastic, obeying the von Mises yield criterion. An upper shear strength limit of $4\tau_{\mathrm{m}}^{\mathrm{u}}/\pi$ was predicted for $V_{\mathrm{f}} \to 1$, which is a 27% increase on the lower bound value $\tau_{\mathrm{m}}^{\mathrm{u}}$. Hancox (1982) showed that for a well-bonded, carbon fiber-reinforced epoxy material, the predicted and experimental values are in tolerable agreement for $V_{\mathrm{f}} \leqslant 0.6$.

In the case where the fiber–matrix interface is the weakest link, it has been suggested (Purslow 1976) that:

$$S_{12} = \frac{\tau_{\mathrm{m}}^{\mathrm{u}}}{(1 + V_{\mathrm{f}})^2}. \tag{21}$$

This approach ignores the presence of stress concentrations within the lamina, which increase in magnitude as $V_{\mathrm{f}} \to 1$. It is possible to obtain reasonable agreement between the experimental data and the following empirical relationship (Chamis 1987, 1989):

$$S_{12} = [1 - (\sqrt{V_{\mathrm{f}}} - V_{\mathrm{f}})(1 - G_{\mathrm{m}}/G_{12\mathrm{f}})]\tau_{\mathrm{m}}^{\mathrm{u}} \tag{22}$$

For laminae with a brittle matrix, the stress concentration effect will cause the strength of the composite to be lower than that of the unreinforced matrix material. If the matrix is able to relieve these stress concentrations through local deformation processes, then $S_{12} \approx \tau_{\mathrm{m}}^{\mathrm{u}}$.

4. *Conclusion*

The current state of the art for modeling the thermoelastic properties of undamaged, unidirectional composites from fiber and matrix properties is well developed and used in design, and it is unlikely to be improved significantly in the future. For regular, hexagonal arrays of fibers the transverse isotropic properties are well modeled by analytical formulae (McCartney 1992a), as has recently been demonstrated for glass- and fiber-reinforced plastics, and for silicon carbide-reinforced aluminium alloy composites, by comparing predictions with FEA solutions and other analytical techniques (unpublished work by Li and McCartney 2000). Issues of future interest are the effects of nonuniform fiber distributions on the properties of both undamaged and damaged composites.

The failure of most unidirectional fiber-reinforced composites, tested in the longitudinal direction, is dominated by the behavior of the fibers whose strength is statistically distributed. While such composites are rarely used in practice, they are however used in fiber-reinforced laminates, whose failure is often dominated by the behavior of the fibers in the 0° plies of a laminate. The failure mechanisms and statistical behavior of UD composites are thus highly relevant to laminate behavior, but proper account must be taken of the triaxial stress states arising in the individual plies of a laminate and of the effects of thermal residual stresses arising from thermal expansion mismatch between the plies.

In current design methodologies, the failure of individual fibers is rarely considered. The conventional approach is to apply one or more of a wide variety of laminate-failure criteria to local stress states in laminated structures. Usually no account is taken of the modes of failure that may arise, or of any damage formation that can lead to localized strain softening and to significant reductions in stress concentrations. This is a very important future development area for the design of composite structures and the failure behavior of individual fibers in such structures will be a key factor in the development of physically based models of damage development and failure of composite structures. The major technical requirement is the need for reliable scale bridging of composite performance from the fiber–matrix level to the ply level, and then to the laminate level, taking full account of complex loading modes and damage formation and the various modes of failure that can arise.

Bibliography

Aveston J A, Cooper G A 1973 Theory of multiple fracture of fibrous composites. *J. Mater. Sci.* **8**, 352–62

Aveston J A , Cooper G A, Kelly A 1971 Single and multiple fracture. In: *Conf. Proc. Properties of Fiber Composites.* IPC Science & Technology, Guildford, UK, pp. 15–26

Berbinau P, Soutis C, Guz I A 1999 *Compos. Sci. Technol.* **59**, 1451–5

Beyerlein I J, Phoenix S L 1996 Stress concentrations around multiple fiber breaks in an elastic matrix with local yielding or debonding using quadratic influence superposition. *J. Mech. Phys. Solids* **44** (12), 1997–2039

Budiansky B 1983 Micromechanics. *Comput. Struct.* **16** (1–4), 3–12

Budiansky B, Hutchinson J W, Evans A G 1986 Matrix fracture in fiber-reinforced ceramics. *J. Mech. Phys. Solids* **34**, 167–89

Chamis C C 1987 Simplified composite micromechanics for predicting microstresses. *J. Reinforced Plast. Compos.* **6**, 268–9

Chamis C C 1989 Mechanics of composite materials: Past, present and future. *J. Compos. Technol. Res.* **11** (1), 3–14

Christensen R M, Lo K H 1979 Solutions for effective shear properties in three-phase sphere and cylinder models. *J. Mech. Phys. Solids* **27**, 315–30

Clyne T W, Withers P J 1993 *An Introduction to Metal Matrix Composites.* Cambridge University Press, Cambridge pp. 12–165

Coleman B D 1958 On the strength of classical fibers and fiber bundles. *J. Mech. Phys. Solids* **7**, 60–70

Cox B N 1991 Extrinsic factors in the mechanics of bridged cracks. *Acta Metall. Mater.* **39**, 1189–202

Cox B N, Marshall D B 1991 Stable and unstable solutions for bridged cracks in various specimens. *Acta Metall. Mater.* **39**, 579–89

Cox B N, Rose L R F 1994 Time- or cycle-dependent crack bridging. *Mech. Mater.* **19**, 39–57

Cox H L 1952 The elasticity and strength of paper and other fibrous materials. *Br. J. Appl. Phys.* **3**, 72–9

Curtin W A, Ahn B K, Takeda N 1998 Modelling brittle and tough stress-strain behaviour in unidirectional ceramic matrix composites. *Acta Mater.* **46**, 3409–20

Curtin W A, Takeda N 1998a Tensile strength of fiber-reinforced composites: I. Model and effects of local fiber geometry. *J. Compos. Mater.* **32**, 2042–59

Curtin W A, Takeda N 1998b Tensile strength of fiber-reinforced composites: II. Application to polymer composites. *J. Compos. Mater.* **32**, 2060–81

Curtis P T 1986 A computer model of the tensile failure process in unidirectional fiber composites. *Compos. Sci. Technol.* **27**, 63–86

Daniels H E 1945 The statistical theory of the strength of bundles of threads. *Proc R. Soc. London, Ser. A* **183**, 405–35

Dvorak G J 1991 Plasticity theories for fibrous composite materials. In: Everett R K , Arsenault R J (eds.) *Metal Matrix Composites: Mechanisms and Properties.* Academic Press, Boston, pp. 1–78

Foster G C, Ibnabdeljalil M, Curtin W A 1998 Tensile strength of titanium matrix composites: direct numerical simulations and analytic models. *Int. J. Solids Struct.* **35** (19), 2523–6

Gumbel E J 1954 Statistical theory of extreme values and some practical applications. Applied Mathematics Series, **33** National Bureau of Standards, Washington, DC

Hancox N L 1982 Shear Properties of Laminates. In: Pritchard G (ed.) *Developments in Reinforced Plastics.* Applied Science, London, UK, Vol. 2, pp. 109–30

Hashin Z 1983 Analysis of composite materials—a survey. *J. Appl. Mech.* **50**, 481–505

Hill R 1964 Theory of mechanical properties of fiber-strengthened materials: II. Inelastic behaviour. *J. Mech. Phys. Solids* **12**, 213–8

Hutchinson J W, Jensen H M 1990 Models of fiber debonding and pull-out in brittle composites with friction. *Mech. Mater.* **9**, 139–63

Iyengar N, Curtin W A 1997 Time-dependent failure in fiber-reinforced composites by matrix and interface shear creep. *Acta Mater.* **45** (8), 3419–29

Marshall D B, Cox B N 1987 Tensile fracture of brittle matrix composites: influence of fiber strength. *Acta Metall.* **35**, 2607–19

Marshall D B, Cox B N, Evans A G 1985 The mechanics of matrix cracking in brittle matrix fiber composites. *Acta Metall.* **30**, 2013–21

McCartney L N 1987 Mechanics of matrix cracking in brittle matrix fiber-reinforced composites. *Proc. R. Soc. London, Ser. A* **409**, 329–50

McCartney L.N. 1989 New theoretical model of stress transfer between fiber and matrix in a uniaxially fiber-reinforced composite. *Proc. R. Soc. London, Ser. A* 425, 215–44

McCartney L.N. 1992a Analytical method of calculating the thermo-elastic constants of a multiphase unidirectional composite. NPL Report DMM(A)57. National Physical Laboratory, Teddington, UK

McCartney L N 1992b Mechanics for the growth of bridged cracks in composite materials. I: Basic principles. *Compos. Technol. Res.* **14**, 133–46

McCartney L N 1992c Mechanics for the growth of bridged cracks in composite materials. II. Applications. *Compos. Technol. Res.* **14**, 147–54

McCartney L N 1996 Stress transfer mechanics: models that should be the basis of life prediction methodology. *Proc. ASTM Symp. Life Prediction Methodology for Titanium*

Matrix Composites (March 1994). ASTM, Philadelphia, pp. 85–113

McCartney L.N. 1999a Analytical model for debonded interfaces associated with fiber fractures or matrix cracks. *Proc. ICCM-12*

McCartney L.N. 1999b Simulation of progressive fiber failure during the tensile loading of unidirectional composites. *NPL Report* CMMT(A)212. National Physical Laboratory, Teddington, UK

McCartney L N, Smith R L 1983 Statistical theory of the strength of fiber bundles. *J. Appl. Mech.* **50**, 601–8

McMeeking R M, Evans A G 1990 Matrix fatigue cracking in fiber composites. *Mech. Mater.* **9**, 217–27

Nairn J A 1992 A variational mechanics analysis of the stresses around breaks in embedded fibers. *Mech. Mater.* **13**, 131–54

Nayfeh A H 1977 Thermomechanically induced interfacial stresses in fibrous composites. *Fiber Sci. Technol.* **10**, 195–209

Pedersen O B 1990 Thermoelasticity and plasticity of composites. II: A model system. *Acta Metall.* **38**, 1201–19

Phoenix S L, Smith R L 1983 A comparison of probabilistic techniques for the strength of fibrous materials under local load-sharing among fibers. *Int. J. Solids Struct.* **19**, 479–96

Pochiraju K V, Tandon G P, Pagano N J 1999 Analyses of single-fiber pushout considering interfacial friction and adhesion. *Composites B: Engineering* (submitted)

Purslow D 1976 *The Shear of Unidirectional Carbon Fibre Reinforced Plastics and their Experimental Determination*. MOD CP No. 1381 Ministry of Defence, London

Rosen W B, Hashin Z 1970 Effective thermal expansion coefficients and specific heats of composite materials. *Int. J. Eng. Sci.* **8**, 157–73

Shu L S, Rosen B W 1967 Strength of fiber-reinforced composites by limit analysis methods. *J. Compos. Mater.* **1**, 366–81

Sideridis E 1988 The in-plane shear modulus of fibre reinforced composites as defined by the concept of interphase. *Compos. Sci. Technol.* **31**, 35–53

Skudra A M 1985 Micromechanics of Failure of Reinforced Plastics. In: Sih G C, Skudra A M (eds.) *Handbook of Composites*. Elsevier, Amsterdam, Vol. 3

Smith R L 1982 The asymptotic distribution of the strength of a series parallel system with equal load-sharing. *Ann. Probability* **10**, 137–71

Smith R L, Phoenix S L 1981 Asymptotic distributions for the failure of fibrous materials under series-parallel structure and equal load sharing. *J. Appl. Mech.* **48**, 75–82

Smith R L, Phoenix S L, Greenfield M R, Henstenburg R B, Pitt R E 1983 Lower tail approximations for the probability of failure of three-dimensional fibrous composites with hexagonal geometry. *Proc. R. Soc. London, Ser. A* **388**, 353–91

Soutis C 1991 Measurement of the static compressive strength of carbon-fibre/epoxy laminates. *Compos. Sci. Technol.* **42**, 373–92

Suresh S, Brockenbrough J R 1993 Continuum models for deformation: metals reinforced with continuous fibers. In: Suresh S, Mortensen A, Needleman A (eds.) *Fundamentals of Metal Matrix Composites*. Butterworths, Boston, pp. 139–57

Tandon G P, Pagano N J 1996 Matrix-crack impinging on a frictional interface in unidirectional brittle matrix composites. *Int. J. Solids Struct.* **33**, 4309–26

Tandon G P, Pagano N J 1998 Micromechanical analysis of the fiber push-out and re-push test. *Compos. Sci. Technol.* **58**, 1709–25

Weaver C W, Williams J G 1975 *J. Mater. Sci.* **10**, 1323–33

Weibull W 1951 A statistical distribution function of wide applicability. *J. Appl. Mech.* **18**, 293–7

L. N. McCartney and W. R. Broughton

National Physical Laboratory, Teddington, UK

Continuous Parallel Fiber Composites: Fracture

The purpose of this article is to provide a brief overview of the micromechanics associated with fracture of unidirectional composites with particular emphasis on toughening mechanisms. Overload fracture is a common failure mode for high-strength, high-stiffness materials. Crack-like flaws are inevitably generated in a material either during fabrication or during service life. If the severity of the flaw is large or the load is high enough, then the flaw becomes the initiation site for catastrophic unstable fracture. A major advantage of composites, however, is that their microstructure promotes the accumulation of small flaws (damage), rather than the growth of a dominant crack. This mechanism of damage accumulation imparts significant energy dissipation capability or fracture toughness to the material, and is most exploited in brittle matrix composite systems.

Composites are most often classified in terms of their matrix, and are designated as polymer matrix composites (PMCs), metal matrix composites (MMCs), or ceramic matrix composites (CMCs). Although these systems possess widely different mechanical properties, they experience similar damage accumulation processes. More significantly, although damage introduces a high level of complexity in determining the stress field ahead of a crack tip or a notch, energy-based fracture mechanics concepts allow an elegant means of characterizing the condition for failure, which spans across all the three matrix systems. These issues on the mechanics and mechanisms of fracture of unidirectional composites are discussed in this article.

As an introduction to the fracture of a material from a dominant flaw, it is useful to consider an elliptic hole embedded in an isotropic or an orthotropic material under plane strain deformation (Fig. 1). The major and minor axis are represented by a and b, respectively, so that the radius of curvature at the end of the major axis (ρ) is b^2/a. For a far-field applied stress, σ_a, perpendicular to the major axis, the maximum stress σ_{yy} in the loading direction occurs at the crack tip (A), and is given by:

$$\sigma_{yy} = \sigma_a\left(1 + 2\sqrt{\frac{a}{\rho}}\right) \tag{1}$$

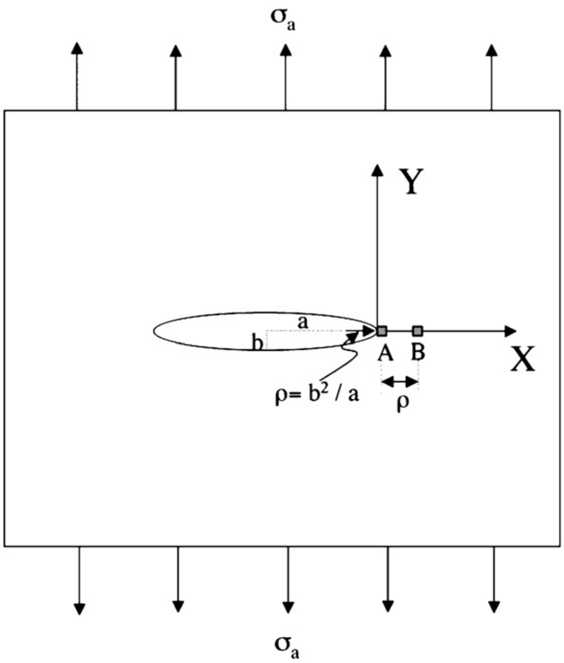


Figure 1
Sketch illustrating an infinite specimen with a notch.

for an isotropic material (Timoshenko and Goodier 1951, Inglis 1913), and by:

$$\sigma_{yy} = \sigma_{\mathrm{a}} \left\{ 1 + \sqrt{\frac{a}{\rho}} \left[2 \left(\sqrt{\frac{E_{yy}}{E_{xx}}} - \nu_{yx} \right) + \frac{E_{yy}}{G_{xy}} \right]^{1/2} \right\} \qquad (2)$$

for an orthotropic material (Bishop 1972, Green and Taylor 1945), when loaded in one of the material's principal directions. Here E_{xx} and E_{yy} are the elastic moduli in the principal x and y directions, respectively, and G_{xy} is the in-plane shear modulus of the composite. Also, ν_{yx} is the principal Poisson's ratio, equal to the normal contraction in the x direction divided by the tensile elongation in the y direction, when a uniaxial stress is applied in the y direction.

There are two important implications of these results:

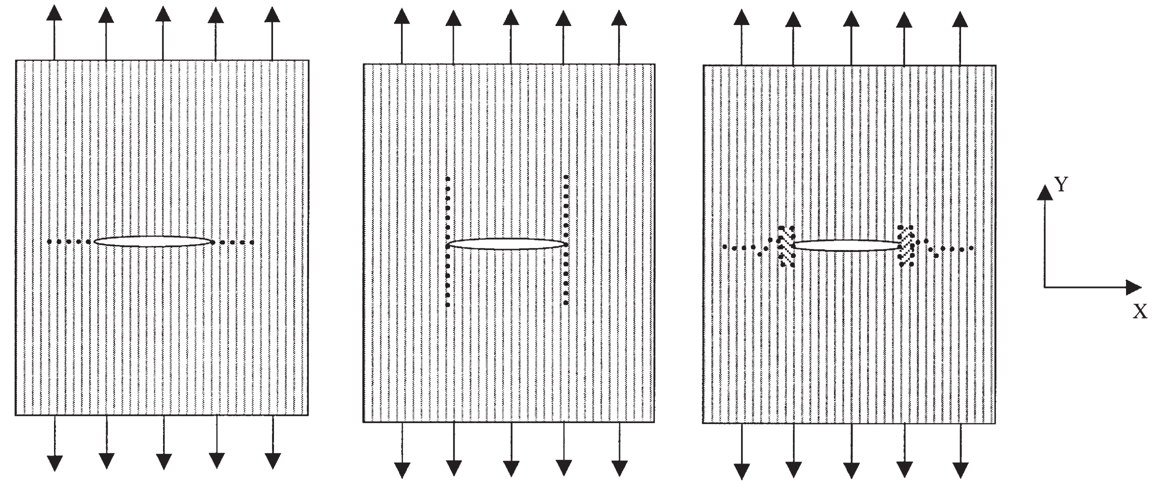
(i) In the limit for a sharp crack, when b and ρ approach zero, the stress σ_{yy} at A becomes singular (tends towards infinity). Under this condition, the simple stress-based design concept of a stress concentration factor (SCF) to specify the condition for failure at the crack tip can no longer be used.

(ii) At any value of ρ, the σ_{yy} stress at the notch tip is significantly higher (typically up to 3.5 fold) in a longitudinally loaded unidirectional lamina of comparatively high E_{yy} than in an isotropic material.

Item (i) requires the concepts of fracture mechanics and is treated in greater detail later. However, a few comments regarding item (ii) are worth mentioning here. The higher σ_{yy} stress for a longitudinally loaded unidirectional lamina may imply that it is more susceptible to transverse fracture (fracture plane perpendicular to the fiber axis) than an isotropic material, for the same geometry and far-field applied stress. However, the anisotropic properties of a unidirectional lamina, primarily the weak strengths in the transverse direction and in shear, introduce crack blunting and stress–relaxation pathways that reduce the level of σ_{yy} ahead of the notch tip. These stress-reduction or crack-shielding mechanisms often make the composite more fracture resistant than the monolithic constituents. Interface mechanics and mechanisms play a critical role in determining the damage tolerance of a composite structure, and are discussed in *Composites: Interfaces** (Cook and Gordon 1964, Kendall 1975, Warrier *et al.* 1997, Gupta *et al.* 1992, He and Hutchinson 1989, Majumdar 1998).

In the case of fracture of a unidirectional lamina with a crack perpendicular to the fiber axis, the unit energy dissipation processes include crack blunting by fiber-matrix delamination, crack bridging by the intact fibers, fiber fracture, matrix plasticity, and crack pinning or trapping. The first tier composite properties, such as the elastic modulus and tensile strength in the principal material directions, dictate how the composite might respond to the presence of a sharp crack. In one case (see Fig. 2(a)) damage may accrue ahead of the notch and lead to crack growth along the original notch/crack plane. Alternatively, if the transverse strength is very low compared to the longitudinal strength, large-scale blunting by splitting along the fiber axis may occur (Fig. 2(b)). The principal stress in the loading direction (σ_{yy}) is thereby relaxed ahead of the crack tip, making the material highly resistant to the propagation of a dominant crack. A low shear strength can also produce a similar blunting effect and contribute to the material's resistance to crack growth. Indeed, some systematic experiments performed by Cooper and Kelly (1967) indicated that if the lamina shear strength was less than 8% of the longitudinal strength, then large-scale shear band growth parallel to the fibers occurred in preference to crack propagation along the original plane of the notch. A more general fracture mode is a combination of damage ahead of the notch in both the x and y directions (Fig. 2(c)), followed by tortuous crack growth in an overall direction perpendicular to the loading axis.

In monolithic materials, the crack path morphology follows that of Fig. 2(a). Ahead of the crack tip, the different stress components then bear a constant ratio to one another as the crack grows, although their overall magnitude increase with crack length; note from Eqn. (1) that the stress σ_{yy} ahead of the crack tip at constant σ_{a} is proportional to $\sqrt{a}$. This type of crack growth is termed as *self-similar* growth. Alternatively, the crack growth profile in



(a) (b) (c)

Figure 2
Failure modes in a longitudinally loaded notched/cracked lamina, with the dots indicating the crack extension paths. (a) Crack extension along the original crack plane; (b) damage and crack growth parallel to the loading direction; and (c) damage along the loading direction, followed by tortuous crack growth in an overall direction perpendicular to the loading axis.

most composites follows Figs. 2(b) and 2(c), where the stress components no longer bear the same ratio with growth of the crack. This loss of self-similarity makes the application of established rules of fracture mechanics rather questionable. However, experimental data indicate that for a given mode of fracture (Fig. 2), the energy required for a significant advance of the initial crack is often characteristic of the material, being only slightly dependent of the specimen geometry. These observations have provided impetus in extending the energy concepts of fracture mechanics, developed primarily for isotropic materials, to composites, thereby allowing a means to assess damage tolerance of a structure in the presence of a notch or a crack-like flaw.

1. Fracture Mechanics Concepts

Fracture of materials can be classified under three major modes of loading, Mode I (or opening mode), Mode II (or in-plane shear mode), and Mode III (or out-of-plane shear mode). Confining attention initially on an isotropic elastic material, the stresses ahead of the crack tip for any combined mode of loading can be expressed as:

$$\sigma_{ij} = \frac{K_{\mathrm{I}}}{\sqrt{2\pi r}} f_{ij}^{\mathrm{I}}(\theta) + \frac{K_{\mathrm{II}}}{\sqrt{2\pi r}} f_{ij}^{\mathrm{II}}(\theta) + \frac{K_{\mathrm{III}}}{\sqrt{2\pi r}} f_{ij}^{\mathrm{III}}(\theta) \quad (3)$$

where r is the radial distance from the crack tip, the $f_{ij}(\theta)$ functions provide an angular dependence, and K_{I}, K_{II}, and K_{III} are the stress intensity factors that characterize the magnitude of the stress singularity for Modes I, II, and III type of loading, respectively (Paris and Sih 1964, Tada *et al.* 1973). In the context of Fig. 1 and Eqn. (1), K_{I} can immediately be identified as $\sigma_{\mathrm{a}}\ (\pi a)^{1/2}$. The stress intensity factors are related to the elastic energy released per unit advance of the crack tip (designated by the symbol, G) through equations of the form:

$$G_{\mathrm{I}} = \frac{(1 - \nu^2)K_{\mathrm{I}}^2}{E} \quad (4)$$

where Eqn. (4) is for a plane strain case and for a pure Mode I type of loading.

Note that K or G characterize the driving force due to an applied load, and can be calculated for any given crack length and load using analytical or computational procedures. Fracture under Mode I loading is postulated to occur when K_{I} or G_{I} reaches a critical value, designated as K_{Ic} and G_{Ic}, respectively, known as the *fracture toughness* of the material. The applicability of this concept in monolithic materials has been validated experimentally using specimens with widely different geometry and crack length. Similar arguments hold good for combined modes of loading, and the effective G is given by:

$$G_{\mathrm{total}} = G_{\mathrm{I}} + G_{\mathrm{II}} + G_{\mathrm{III}} \quad (5)$$

Although K_{Ic} or G_{Ic} can be interchangeably used to characterize the criticality of fracture, they have traditionally involved two modes of thinking. In the case of K, failure is viewed as occurring when a critical stress field is realized ahead of the crack tip. Alternatively, G is most often viewed in terms of the overall energetics of failure, making it more amenable to experimental determination. Thus, it can be shown that G can simply be expressed as:

$$G = \frac{P^2}{2B}\frac{\mathrm{d}C}{\mathrm{d}a} \tag{6}$$

where P is the applied load, C is the specimen compliance, which is a function of crack length a, and B is the specimen thickness.

The lack of self similarity of crack growth in composites makes it difficult to ascribe the onset of crack advance in terms of a defined level of stress field. Alternatively, G_{Ic} retains its significance, in that it still characterizes the level of elastic energy released per unit of crack advance. That is why the energy release rate parameter is most often used in the composites literature, although the K_{Ic} level is also specified through equations of the form shown in Eqn. (5).

There are two additional parameters that become important when inelastic processes dominate. One is the J integral, which is a path-independent integral for any path around the crack tip, and represents the energy associated with a unit advance of the crack tip under situations where stress increases monotonically. Thus, J is quantitatively equal to and synonymous with G for initiation of crack propagation; however, the J integral more rigorously accounts for inelastic processes near the crack tip, and in monolithic materials it characterizes the crack-tip stress and strain fields, similar to K in linear elastic solids.

A final parameter that often enters into fracture mechanics formulations is the crack tip opening displacement (δ), which is the total displacement at the crack tip due to elastic and inelastic deformation processes. In the case of ductile metal alloys, the largest contribution to δ comes from plasticity, and at the onset of fracture it is of the order of distance between precipitates/second phase particles where ductile void nucleation occurs. When plasticity is the dominant inelastic mechanism, δ is related to J through equations of the form:

$$\delta = \frac{kJ}{\sigma_{\mathrm{o}}} \tag{7}$$

where k is of the order of unity, and σ_{o} is the yield strength of the material in uniaxial tension. At the onset of fracture, δ attains a critical value δ_{c} associated with a critical value of J_{Ic} (or equivalently G_{Ic} or K_{Ic})

2. *Fracture Mechanisms*

Physically δ_{c} represents the blunting of the crack tip due to inelastic deformation. In monolithic materials it ranges from submicron to about 10–20 μm. Alternatively, the crack opening displacements in composites are much larger due to damage ahead of the crack tip, making the use of conventional fracture mechanics difficult. Unlike highly plastic solids, where the entire structure may behave nonlinearly due to spread of the yield zone, in composites the material behaves elastically in most of the structure, and large nonlinear processes dominate only in the near crack tip region. Therefore, it is convenient to describe the toughness of the composite in terms of a local crack separation law, which now includes all nonlinear effects.

Figure 3 illustrates such a traction law, and the fracture resistance of different composites (PMCs, MMCs, and CMCs) can be traced to variations in this traction law due to different inelastic mechanisms. The area under the curve represents the work of fracture contributed by such a local traction law, and is best designated by its contribution to the energy spent in propagation the crack in the form of parameter ΔG:

$$\Delta G = \int_0^{\delta_{\mathrm{c}}} \sigma(\delta)\mathrm{d}\delta \tag{8}$$

In Fig. 3, the rising portion of the curve is due to increasing stresses that take the material point at the crack tip from an initial state to a state of maximum tension. The declining portion is due to different separation mechanisms, such as debonding, pull-out, etc. Most of the analytical formulations in the fracture mechanics of composites deal with the rising portion of the traction separation law, implicitly assuming that there is a vertical drop off to zero load from the point of maximum stress. However, the declining portion can contribute substantially to the work of fracture.

In this section, attention will be concentrated on cracked unidirectional composites loaded in the fiber

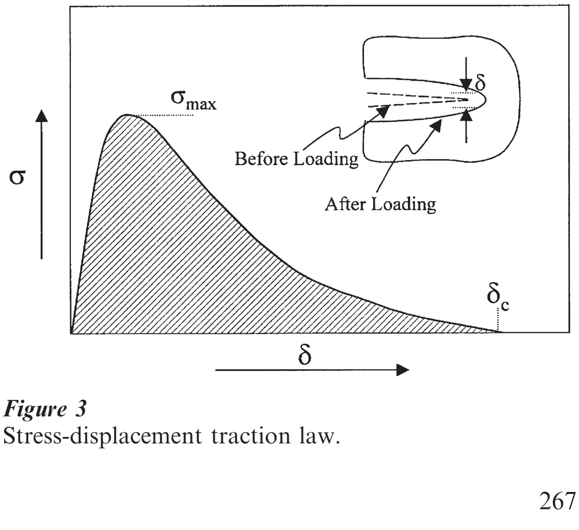


Figure 3
Stress-displacement traction law.

direction (Figs. 2(a)–2(c)), since the best properties of the composite are obtained in this configuration. Also, at this point it is important to distinguish between different types of systems: (i) brittle matrix system, and (ii) ductile matrix system. The former includes CMCs and PMCs based on brittle polymers like polyimides, while the latter includes MMCs and PMCs that are made with ductile polymers like thermoplastics and toughened resins. This distinction is needed because failure of composites with brittle matrices is preceded by transverse matrix cracks that propagate from the crack/notch tip, leaving intact fibers in their wake. Alternatively, in ductile-matrix systems, it is the fibers that crack first which are then followed by bridging by the ductile matrix. In CMCs, damage tolerance is achieved through a weak fiber-matrix interface. Otherwise, matrix cracks can propagate right through the fiber, making the brittle matrix composite of little engineering value. Energy dissipation in these systems is largely through: (i) crack bridging by the intact fibers, which produce a stress intensity factor (K_B) that opposes the K_{appl} due to the applied load (this is also called crack tip shielding), and (ii) pull-out of cracked fibers, where the fiber cracks are located at some nonzero distance from the crack-plane. In ductile matrix systems, energy dissipation is completely dominated by plastic deformation of the ductile matrix between the cracked fibers. One consequence of this difference is that a higher fiber volume fraction often enhances toughness in brittle matrix systems, whereas it can reduce toughness in ductile matrix systems.

2.1 *Brittle Matrix Systems*

In terms of the mechanisms of fracture, it is useful to consider the two configurations depicted in Figs. 4(a) and 4(b). These are designated as Case I and Case II, respectively.

In Fig. 4(a), a steady-state condition is considered wherein a single matrix crack emanates from the notch tip at a sufficiently low load. As this matrix crack grows, leaving behind intact fibers, a situation is reached where the fiber at the notch tip just begins to fail. Thereafter, failure of the composite can occur in a steady state manner, and the fracture toughness corresponds to the applied stress intensity factor at which lead fiber fails. The toughening associated with Case I has been worked out by Budiansky and Amazigo (1989). In Case II, we consider fracture of fibers that have a statistical strength distribution so they fail at different distances from the crack plane. Final failure requires that these fibers pull-out of the matrix. Because of frictional resistance, substantial energy is dissipated in the decreasing segment of the stress-displacement traction law (Fig. 3), contributing to significant increase in the toughness of the composite. The work of pull-out has been worked out by a number of investigators (Kelly 1970, Phillips 1972),

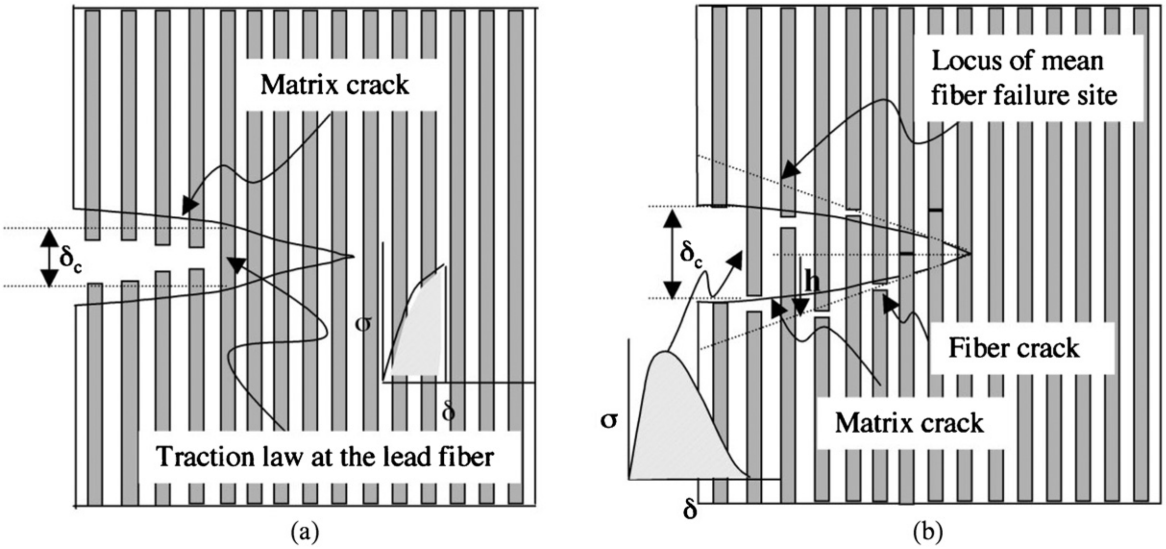

Matrix crack
δc
σ
δ
Traction law at the lead fiber
(a)
Locus of mean fiber failure site
δc
h
Fiber crack
Matrix crack
σ
δ
(b)

Figure 4
Crack propagation in brittle matrix composites. (a) Steady-state bridging of matrix cracks by intact fibers, with the lead fiber failing at the original crack plane. (b) Matrix crack bridging, but in this case the fibers fail statistically at different distances from the crack plane. The traction law is sketched for the two cases.

and we present here the results of Thouless and Evans (1988).

(*a*) *Case I.* As shown in Budiansky and Amazigo (1989), a simple J integral calculation provides the following relationship:

$$\frac{J_{\text{applied}}}{A} = (1-c)J_{\text{matrix}} + \int_0^{\delta_c} \sigma(\delta)\mathrm{d}\delta \qquad (9)$$

where c is the volume fraction of fibers, A is a factor that accounts for the orthotropy of the sheet in Fig. 4(a) ($=1.0$ for identical elastic constants of the fiber and the matrix), $J_{\text{appl}} = (1-v^2)K^2_{\text{applied}}/E_c$, and $J_{\text{matrix}} = (1-v^2)K^2_m/E_m$, where K_m is the critical stress intensity factor for the monolithic matrix. Also, E_m is the modulus of the matrix, and E_c is the composite longitudinal modulus ($= cE_f + (1-c)E_m$, where E_f is fiber modulus). Values of A as a function of c and E_f/E_m are plotted in Budiansky and Amazigo (1989). Note that J is synonymous with G, but is more appropriate here because of inelastic processes. Also, if there are parallel matrix cracks emanating from the notch tip, then the first term in Eqn. (9) is multiplied by the number of those matrix cracks.

Assuming relative sliding/slipping between the cracked matrix and the intact fibers in the wake of the matrix crack, it follows that:

$$\frac{J_{\text{applied}}}{J_{\text{matrix}}} = A(1-c)\left[1 + \frac{2c\sigma_f^{\text{avg}}}{\sigma_{\text{ss}}}\right] \qquad (10)$$

where σ_f^{avg} is the average bundle strength of the fiber, and σ_{ss} is the steady-state matrix cracking stress (sometimes referred to as the pseudo yield stress) for the composite. As shown in the ACK theory (Aveston, Cooper, and Kelly 1971), and also analyzed in the BHE theory (Budiansky, Hutchinson, and Evans 1986) σ_{ss} can be represented as:

$$\sigma_{\text{ss}} = E_c\left[\frac{6c^2E_f\tau}{(1-c)E_cE_m}\right]^{1/3}\left[\frac{J_{\text{matrix}}}{rE_m}\right]^{1/3} \qquad (11)$$

where τ is the axial frictional sliding stress between the fiber and the matrix, and r is the radius of the fiber. Note that $c\sigma_f^{\text{avg}}$ is approximately the breaking strength of an unnotched longitudinal composite, since only the fibers carry the entire load after matrix cracks run through the entire length of the composite at a stress of σ_{ss}. A typical value of $c\sigma_f^{\text{avg}}/\sigma_{\text{ss}}$ is five, so that Eqn. (10) indicates that a 10-fold toughness increase is obtainable based on this mechanism of toughening.

It is useful to note that the toughness ratio varies inversely as σ_{ss} (Eqn. (10)). Thus, while a higher frictional resistance (τ) and a smaller fiber radius (r) would increase the "yield strength" of the brittle matrix composite, it is detrimental to the fracture toughness. Consequently, material design will need to be optimized, depending upon the desired application in components.

(*b*) *Case II.* In this case, we consider the statistical failure of brittle fibers, whose strength can be expressed by the two-parameter Weibull distribution. In this methodology, the cumulative probability of failure (P_f) of a fiber at a stress σ is given by:

$$P_f = 1 - \exp^{-(L/L_0)(\sigma/\sigma_0)^m} \qquad (12)$$

Here σ_o is the characteristic fiber strength at length L_o, and is usually designated as the Weibull strength, and m is the Weibull modulus.

The problem of Fig. 4(b) was analyzed by Thouless and Evans (1988). Figure 5 from Evans (1990) shows the traction law, in nondimensional co-ordinates, that result for different values of Weibull modulus, m. Here a value of infinity corresponds to a deterministic strength (similar to Case I), and may occur in epoxy based composites toughened with polymeric fibers, or even with SiC fibers ($m \sim 15$ in the virgin condition). Alternatively, a value of $m = 5$ more closely resembles situations that may be encountered in CMCs. Note that as m decreases, the declining portion of the traction law contributes more and more to the toughness of the composite. In terms of the design parameters, the toughness increase is approximately proportional to $\{r^{m-3}/\tau^{m-1}\}^{1/(m+1)}$. Thus, the toughness increases with increasing fiber radius when $m > 5$, but decreases when $m < 3$. Conversely, it increases with increasing τ when $m < 1$, and decreases when $m > 2$. This type of behavior arises because of the competing importance of the contribution to toughness from intact bridging

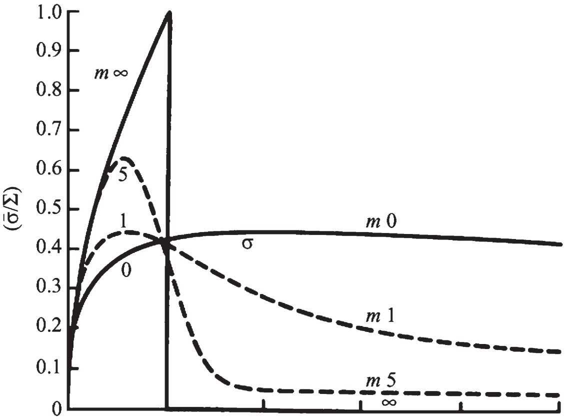


Figure 5
Nondimensional traction law (σ versus u) for the pull-out fiber case, for different values of Weibull modulus, m (after Evans 1990).

fibers and the failed fibers that experience pull out. Thus, knowledge of m and τ are important for optimization of composite toughness under fiber pull-out conditions.

(*c*) *Case of ductile fiber toughening in brittle matrix composites*. This type of toughening mechanism, although developed for particle toughened composites, is also applicable for ductile fiber toughened CMCs and PMCs. Under the condition that the ductile fiber is perfectly plastic and fully yielded ($\sigma = \sigma_Y$), Eqn. (8) shows that the toughening contribution per ductile fiber is simply $k\sigma_Y\delta_c$, where k accounts for a higher strength of the constrained ductile fiber. Assuming no debonding between the cracked matrix and the ductile fiber (for example, a metallic fiber), the value of δ_c is of the order of the fiber diameter, $2r$. The value of k is more difficult to estimate, but should not exceed that ahead of a crack tip, namely 3–4σ_Y. These parameters then allow a fair assessment of toughening associated with ductile fibers in a brittle matrix. A more rigorous treatment (Budiansky *et al.* 1988) leads to the result:

$$\frac{K_{\text{applied}}/K_m}{\sqrt{\bar{\omega}(1-c)}} = \left[1 + \frac{2c}{1-c}\frac{E_m S\delta_c}{K_m^2(1-\nu_m^2)}\right]^{1/2} \qquad (13)$$

where S is the effective strength of the ductile fiber and c is fiber volume fraction. Also,

$$\bar{\omega} = \frac{E_c(1-\nu_m^2)}{E_m(1-\nu_c^2)} \qquad (14)$$

It is useful to note that substantial increase in δ_c is possible if there is relative sliding between the fiber and the matrix, since bridging can be sustained over much larger matrix crack length. Essentially, δ_c becomes the gage length over which the debonding occurs, times the strain to failure in uniaxial tension. However, S is close to σ_Y in this case, so that there is substantial loss of stiffness of the system in the presence of matrix cracks. Since component functionality is also often based on maximum possible displacement, there will always be a need to optimize the interface for component specific performance.

2.2 Ductile Matrix Systems

In ductile matrix systems, a number of different behaviors have been observed depending on the strength of the matrix, the fiber-matrix bond strength, and the volume fraction of fibers.

When the matrix strength is low, matrix dominated shear deformation occurs prior to any fiber fracture. Indeed, in boron fiber reinforced aluminum composites, a deformation mode similar to Fig. 6 is observed (Dvorak *et al.* 1989, Awerbuch and Hahn 1979, Reedy 1980), where an intense slip zone develops over a plastic zone of length (R) perpendicular to the crack plane. At a critical load, when R is on the order of 3–17-fold the crack length, the damage zone stops propagating and is replaced by failure of fibers at the crack tip. This, in turn leads to catastrophic fracture of the composite along the original notch plane. Most notable is the fact that the H-shaped shear zone is created prior to crack extension, and is absent during propagation of the crack, i.e., crack extension is non self-similar. A similar type of damage zone and crack extension has also been observed in glass fiber reinforced epoxy composites (Tirosh 1973).

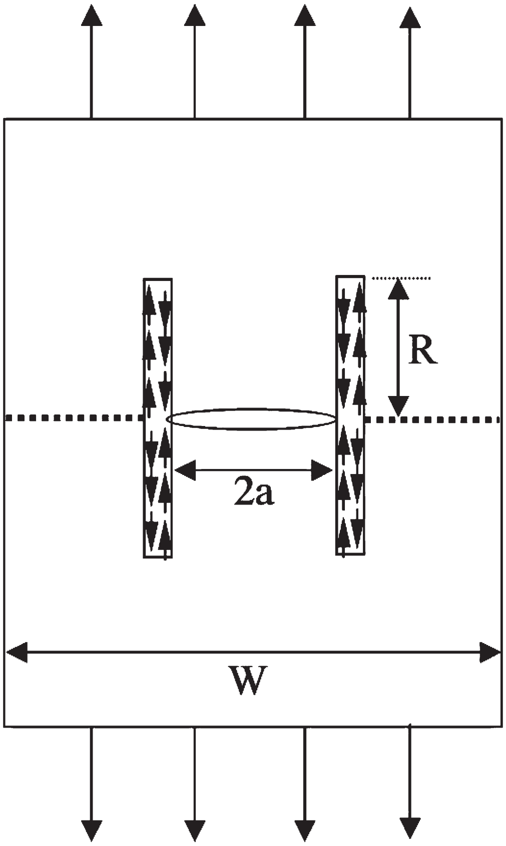


Figure 6
Sketch illustrating H-type plastic zone development at the notch tip, followed by composite fracture along original notch plane.

In the above type of deformation mode, the effective toughening is extremely high, because the blunted crack tip attenuates the stress ahead of the crack tip over a distance of the order of the crack tip opening displacement. For crack lengths of a few millimeters or longer, toughness values of 100 MPa m$^{1/2}$ have been realized for the B/Al system. The effect reduces at smaller crack length, and approaches approximately 30 MPa m$^{1/2}$ at crack length <0.5 mm. Thus, fracture toughness may not be the appropriate parameter for predicting failure in these systems which do not obey self-similar crack growth. Dvorak *et al.* (1989) attempted to estimate the onset of fracture based on attainment of a critical strain in the fiber direction over two fiber diameter distance ahead of the crack tip. It was found that the local strain in the representative volume element for specimens with different notch lengths all fall in the error band for the failure strain of unnotched composites. The following set of equations may be used to estimate the failure load (P_{ult}) for a center-cracked panel of width W and crack length $2a$ (Zarzour and Paul 1992), and possessing an unnotched strength, $\sigma_{\mathrm{unnotched}}$:

$$\frac{P_{\mathrm{ult}}}{\left(1-\dfrac{2a}{W}\right)}+2\frac{\tau^*}{\pi(a_2-a_1)}\left[-a_1^2\ln\left(\frac{1+\sqrt{1+a_1^2\beta^2}}{a_1\beta}\right)+a_2^2\ln\left(\frac{1+\sqrt{1+a_2^2\beta^2}}{a_2\beta}\right)\right]=\sigma_{\mathrm{unnotched}} \qquad (15)$$

where

$$a_{1,2}^2=\left(\frac{E_1}{2G_{12}}-\nu_{12}\right)\pm\sqrt{\left(\frac{E_1}{2G_{12}}-\nu_{12}\right)^2-\frac{E_1}{E_2}} \qquad (16)$$

and

$$b\approx d\left(\frac{2\pi}{3\sqrt{3}c}\right)^{1/2} \qquad \beta=\frac{aP_{\mathrm{ult}}}{b\tau^*}$$

Here, d is the fiber diameter, c is the volume fraction of fibers, τ^* is the in-plane shear strength of the unidirectional composite in the fiber direction, E_1 is the modulus in the fiber direction, E_2 in the transverse direction, G_{12} is the shear modulus, ν_{12} is the Poisson's ratio, and b is of the order of distance between two adjacent fibers.

When the matrix strength is high, or the fiber volume fraction is high, the dominant damage mode is fiber fragmentation in the zone of intense matrix plastic deformation near the crack tip, and ultimately composite fracture. This is illustrated in Fig. 7(a). In order to predict the fracture toughness, some estimate of the flow stress and the critical displacement is needed, for plugging into Eqn. (8). Friler *et al.* (1993) and Argon (2000) modeled the failure process by considering that the periphery of fractured fiber tips act as the nucleation center from which intense matrix plasticity develops (see Fig. 7(b)). This is a form of macroscopic void growth, at a length scale that is significantly larger than the distance between intermetallic particles, which act as the void nucleation site for fracture of monolithic metallic alloys. Using Eqn. (8), Friler *et al.* (1993) estimate the fracture toughness as:

$$J_{\mathrm{Ic}}\approx\frac{1}{\sqrt{3}}(1-c)\sigma_{\mathrm{Y}}L_{\mathrm{D}} \qquad (17)$$

where

$$L_{\mathrm{D}}\approx\beta_{\mathrm{n}}d\,\frac{1-\sqrt{c}}{\sqrt{c}} \qquad (18)$$

and β_{n} is approximately 2. Here d is the diameter of the fiber, c is the volume fraction of fibers, and σ_{Y} is the yield stress of the matrix.

Very few experiments have been conducted on the fracture toughness of titanium matrix composites. Connell *et al.* (1994) reported a toughness of approximately $71\,\mathrm{kJ\,m^{-2}}$ for a SiC/Ti-alloy, which may be compared with a typical toughness of $40\,\mathrm{kJ\,m^{-2}}$ (based on $K_{\mathrm{Ic}}=70\,\mathrm{MPa\,m^{1/2}}$) for monolithic Ti–6Al–4V alloy. Using Eqn. (17) with $\sigma_{\mathrm{Y}}=1040\,\mathrm{MPa}$, $c=0.32$, and $d=100\,\mu\mathrm{m}$, a toughness value of $62\,\mathrm{kJ\,m^{-2}}$ is estimated, which compares reasonably well with the experimental data.

It is useful to note that J_{Ic} predicted by Eqns. (17) and (18) is quite strongly dependent on the volume fraction of fibers. Thus, $(1-c)(1-\sqrt{c}/\sqrt{c})$ decreases from approximately 0.59 to 0.12 on increasing the volume fraction from 0.3 to 0.6. High volume-fraction alumina/aluminum composites are currently being developed for a variety of applications, such as high-tension electrical cables and piston rods. Because of the lower strength of the alloy, and the high volume fraction of alumina fibers, toughness values significantly less than titanium matrix composites are anticipated.

A final note is in order regarding the role of the fiber-matrix interface. Equation (17) shows that the toughness is an increasing function of L_{D}. Weak interfaces would permit greater fiber-matrix sliding, thereby increasing the fracture toughness. This effect was elegantly utilized by Argon *et al.* (1998) to toughen aluminum-based composites, while maintaining acceptable transverse strengths. In SiC reinforced titanium matrix composites, pull-out lengths are typically less than one fiber diameter, even for weak carbon based interfaces. This is largely because of the high radial compressive stress that is generated at the fiber-matrix interface at the tip of a cracked fiber. For strong interfaces, such as those formed without a carbon layer on the SiC monofilaments, the pull-out length will be even shorter. However, the effect of a smaller L_{D} may be balanced by a higher flow stress associated with constrained yielding of the matrix in the fragmentation zone. Tensile tests on unnotched SiC/Ti-matrix composites with different interfaces indicate correlated fiber fractures, independent of the type of interface. Slip band observations and ultrasonic images of fiber breaks confirm that correlated fractures are a result of slip band

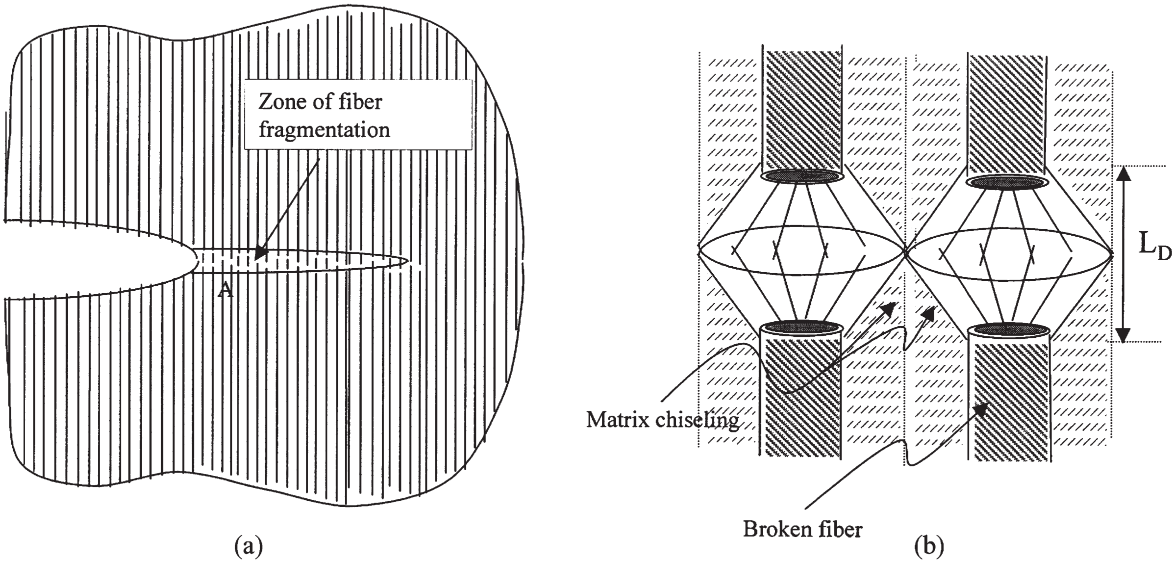


Figure 7
(a) Overall damage pattern in ductile matrix composites. There is a zone of plasticity where fibers are fragmented, followed by propagation of the matrix crack through that damaged region. (b) The Friler model for determining the toughness, assuming chiseling of the matrix from the broken fiber ends (Friler *et al.* 1993).

interactions, whereby a slip band impinging on a fiber localizes sufficient strain to fracture that fiber (Majumdar *et al.* 1998). Similar experiments have to be conducted with notched composites to provide an assessment of the role of interface strength on the toughness of composites with high matrix strength.

2.3 Intermediate Ductility Matrix Systems

An important characteristic of polymer matrix composites is that their behavior can range from brittle to ductile matrix systems depending on the particular polymer used. Moreover, changing the temperature or loading rate (loading history) will alter the ductility for any polymer. Higher temperatures and/or lower loading rates generally increase ductility. Both the amount of ductility and the dependence on temperature and loading history are determined by the microstructure of the polymer. This includes molecular level structure as well as morphology since many polymer matrices are two phase systems where the base polymer contains particles of a polymeric additive to improve toughness and other desirable properties. In a sense, the polymer itself is a composite.

The relationship between microstructure and ductility of polymers has been studied extensively, but is beyond the scope of this article. The reader is referred to some of the many references in this area (Bucknall 1977, Bitner *et al.* 1981, Pearson and Yee 1989, Wu *et al.* 1990, Huang *et al.* 1993, Kinloch *et al.* 1983). What should be noted here, however, is that the effect of polymer ductility on matrix toughness can often be described in terms of a toughening mechanism involving crack tip blunting, see Eqn. (7). Thus, increasing the ductility allows more blunting of the crack tip. This, in turn, reduces the local stress concentration, thereby postponing damage initiation and retarding damage growth. The ultimate result is increased fracture toughness of the polymer. Kinloch *et al.* (1983) have investigated this mechanism, by plotting the toughness of the polymer as a function of crack tip radius, where the radius was obtained either by blunting of an initially sharp crack or by incorporating artificially drilled holes of fine enough radius at the end of a sharp notch. Specimens with sharp cracks were tested at different temperatures and cross-head speeds to vary the amount of blunting, while specimens with the artificial crack tip radii were loaded at high cross-head speeds to minimize further blunting of the crack tip. A clear relationship between the measured toughness and the root radius was found, but more importantly the behavior was independent of how the tip radius was incorporated into the specimen.

In addition to crack blunting, three other toughening mechanisms are also important in polymer composites: crack pinning, crack diversion or bifurcation, and crack bridging. The interaction of these mechanisms can be illustrated by considering the two most common forms of composites fracture in PMCs, i.e., interlaminar and transverse.

(*a*) *Interlaminar fracture.*
Crack Blunting: In interlaminar fracture, the crack propagates in the resin and/or at the resin-fiber interface parallel to the fibers. As a result, the matrix toughness would be expected to play a major role in the interlaminar fracture energy. One example of this is shown in Fig. 8 where the Mode I interlaminar fracture energy is plotted versus the Mode-I fracture toughness of the "neat" resin (i.e., the unreinforced matrix) (Hunston *et al.* 1987). The results show that tougher resins produce tougher composites, but the relationship is not a simple one; see the trend lines in Fig. 8. In the brittle range, the composite is significantly tougher than the resin so the transfer of toughness is good. Beyond a certain point, however, the transfer is less complete. To increase the fracture energy of the composite by $1\,\mathrm{J\,m^{-2}}$, the fracture energy of the resin must be increased by approximately $2\,\mathrm{J\,m^{-2}}$. The proposed explanation is that all crack blunting mechanisms involve inelastic deformation of the matrix material in the crack-tip region, with the toughness increasing with the size of the plastic/viscoplastic zone, consistent with elastic-plastic fracture mechanics principles. For brittle materials, this inelastic zone is smaller than the inter-fiber spacing, so that the resin toughness is completely transferred to the composite. For tougher materials, however, the size of the inelastic zone in the bulk resin is larger than the inter-fiber spacing. Consequently, in the interlaminar fracture experiment, the fibers limit the extent of toughness increase per unit increase in resin toughness. In support of this argument, the estimated size of the deformation zone at the transition point in behavior (intersection of the trend lines in Fig. 8) compare very well with the interfiber spacing between plies.

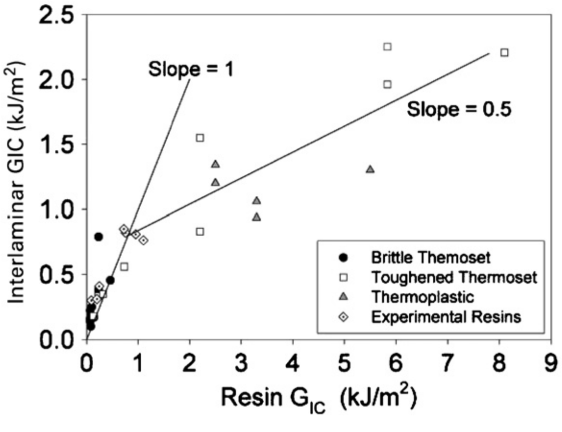


Figure 8
Composite interlaminar fracture energy as a function of the fracture energy for the corresponding resin. Lines indicate the general trends only. The relative standard uncertainty in the determination of G_{Ic} is 15% (after Hunston *et al.* 1987).

Crack Pinning or Trapping: Because interlaminar fracture is an important failure mode in polymeric systems, methods have been developed to improve the fracture toughness. One approach is to add rubber particles between the fiber layers during fabrication. The particles are thereby trapped between the fiber layers and form a barrier to crack propagation. When an interlaminar crack propagates it runs into these barriers and is held back because the particles have a higher fracture toughness than the matrix. In effect, the matrix crack is pinned or trapped by these barriers. Higher loads are then required to force the crack through or around the particles. This toughening mechanism has been shown to be effective even for impact loading indicating that the pinning mechanism is relatively independent of loading rate. This is in contrast to the case of the crack blunting mechanism, where the toughness is significantly rate sensitive. Bower and Ortiz (1990) have modeled the crack trapping behavior and have shown good correlation with experimental results.
Crack Diversion or Bifurcation: It was proposed in some early studies of interlaminar fracture that under the proper conditions, poor fiber-matrix bonding could increase the fracture energy (McKenna *et al.* 1974). For this to occur, the weak interface bonding needed to be localized, randomly distributed along the fibers, and interspersed with areas of good bonding. Under these conditions, a growing crack would be diverted back and forth between the many weakly bonded areas so that the trajectory became very tortuous. Moreover, the crack may be split into several branches that grow for a short distance and then end. When poor interface bonding becomes too extensive, however, the result is a decrease in fracture toughness because the weak interface provides an easy path for crack growth. This is illustrated in Fig. 8 where the thermoplastic resins generally show lower interlaminar fracture toughness than thermoset resins with the same resin fracture energies. Examination of the fracture surface shows that the fiber-matrix interface bonding in the thermoplastics was much poorer than it was in the corresponding thermoset.
Crack Bridging: Another interesting feature in Fig. 8 is that the interlaminar fracture energy for composites with the more brittle resins is almost always higher than the fracture energy for the resin itself. This has been attributed to the fact that the composite has an additional toughening mechanism: fiber bridging (Hunston *et al.* 1987), discussed earlier in this article.

(*b*) *Longitudinal fracture*. These same four types of toughening mechanisms are present when a crack propagates perpendicular to the fibers. As discussed previously, toughening involves a combination of work associated with interface debonding, the work of stretching and breaking fibers, and the work required for fiber pull-out.

With polymers, experiments generally show that the highest resistance to crack growth occurs at intermediate bond strength. Moreover, with very strong interface bonding, a post failure examination shows relatively few fiber ends protruding from the fracture surface. As the interface bond strength decreases, there is an increase in the number of protruding fiber ends and their length. The authors (Kelly 1970, Harris *et al.* 1975) explain these results by noting that when the interface bond is strong, there is resistance to crack diversion and fiber-matrix debonding, and the matrix crack can propagate into the fiber. Consequently, fewer fibers protrude on the fracture surface. As the interface strength is decreased, the amount of crack diversion and interface debonding increases, so that fiber fractures can now occur statistically at locations away from the crack plane. Consequently, there is an increase in energy dissipation due to both debonding and pull-out, consistent with Figs. 3 and 5. However, when the interface strength becomes very low, the work of debonding may actually decrease even with a greater length of debonding, because of lower frictional resistance per unit area of the fiber-matrix interface. For polymer matrix systems, the frictional resistance is quite low, indicating that below a certain bond strength the pull-out energy could also decrease with a decrease in interface strength. Consequently, the highest fracture energy is obtained at an intermediate value of interface strength.

3. Concluding Remarks

Although the variety of composite materials is very large, fracture mechanics provides a common approach to analyze failure. Examining the similarities and differences among the various types of composites shows that the toughening mechanisms often extend across a wide range of composites. Consequently, much can be learned from such studies. In order to design composites intelligently, it is important to develop predictive equations based on the observed micromechanisms of fracture. Such models have to be validated experimentally, and this article illustrates a number of examples where good correlations between experiments and model predictions have been obtained. For additional reading, please refer to Kelly and Macmillan (1986).

Bibliography

Argon A S 2000 Fracture: strength and toughness mechanisms. In: Kelly A, Zweben C (eds.) *Comprehensive Composite Materials*. Pergamon, Oxford, Vol. 1, pp. 763–802

Argon A S, Seleznev M L, Shih C F, Liu X H 1998 Role of controlled debonding along fiber/matrix interfaces in the strength and toughness of metal matrix composites. *Int. J. Fract.* **93**, 351–71

Aveston J A, Cooper G A, Kelly A 1971 Single and multiple fracture. In: *Properties of Fibre Composites*. IPC Science and Technology, Guildford, UK, pp. 15–26

Awerbuch J, Hahn H T 1979 Crack tip damage and fracture toughness of boron/aluminum composites. *J. Compos. Mater.* **13**, 82–107

Bishop S M Stresses near an elliptical hole in an orthotropic sheet. Technical Report 72026. Royal Aircraft Establishment, Farnborough, UK

Bitner J L, Rushford J L, Rose W S, Hunston D L, Riew C K 1981 Viscoelastic fracture of structural adhesives. *J. Adhes.* **13**, 3–28

Bower A F, Ortiz M 1990 Solution of 3-dimensional crack problems by finite perturbation methods. *J. Mech. Phys. Solids* **38**, 443–80

Bucknall C B 1977 *Toughened Plastics*. Applied Science, London

Budiansky B, Amazigo J C 1989 Toughening by aligned, frictionally constrained fibers. *J. Mech. Phys. Solids* **37** (1), 93–109

Budiansky B, Amazigo J C, Evans A G 1988 Small-scale crack bridging and the fracture toughness of particulate reinforced ceramics. *J. Mech. Phys. Solids* **36** (2), 167–87

Budiansky B, Hutchinson J W, Evans A G 1986 Matrix fracture in fiber-reinforced ceramics. *J. Mech. Phys. Solids* **34**, 167–89

Connell S J, Zok F W, Du Z Z, Suo Z 1994 On the tensile properties of a fiber reinforced titanium matrix composite—II. Influence of notches and holes. *Acta Metall.* **42** (10), 3451–61

Cook J, Gordon J E 1964 A mechanism for the control of crack propagation in all-brittle systems. *Proc. R. Soc. London* **A282**, 508–20

Cooper G A, Kelly A 1967 Tensile properties of fiber-reinforced metals: fracture mechanics. *J. Mech. Phys. Solids* **15**, 229–79

Dvorak G J, Bahei-El-Din Y A, Bank L C 1989 Fracture of fibrous metal matrix composites—I. *Experimental results. Eng. Fract. Mech.* **34** (1), 87–123

Evans A G 1990 The mechanical behavior of ceramic composites. In: Embury J D, Thompson A W (eds.) *The Modeling of Material Behavior and Its Relation to Design*. Proc. TMMS Symp. TMMS, Warrendale, pp. 245–67

Friler J B, Argon A S, Cornie J A 1993 Strength and toughness of carbon fiber reinforced aluminum matrix composites. *Mater. Sci. Eng.* **A162**, 143–52

Green A E, Taylor G I 1945 *Proc. R. Soc. London* **A184**, 181

Gupta V, Argon A S, Suo Z 1992 Crack deflection at an interface between two orthotropic media. *J. Appl. Mech., Trans. ASME* **59**, S79–87

Harris B, Morley J, Phillips D C 1975 Fracture mechanisms in glass-reinforced plastics. *J. Mater. Sci.* **10**, 2050–61

He M Y, Hutchinson J W 1989 Crack deflection at an interface between dissimilar elastic materials. *Int. J. Solids Struct.* **25**, 1053–67

Huang Y, Hunston D L, Kinloch A J, Riew C K 1993 Mechanisms of toughening thermoset resins. In: Riew C K, Kinloch A J (eds.) *Toughened Plastics I*. Advances in Chemistry Series 233. American Chemical Society, Washington, DC, pp. 1–35

Hunston D L, Moulton R J, Johnston N J, Bascom W D 1987 In: Johnston N.J. *Toughened Composites*, ASTM STP 937. American Society for Testing and Materials, Philadelphia, pp. 74–94

Inglis C E 1913 *Trans. Inst. Nav. Archit.* **55**, 219

Kelly A 1970 Interface effects and the work of fracture of a fibrous composite. *Proc. R. Soc. London* **A319**, 95–116

Kelly A, Macmillan N H 1986 *Strong Solids*, 3rd Edn. Clarendon, Oxford

Kendall K 1975 Transition between cohesive and interfacial failure in a laminate. *Proc. R. Soc. London* **A344**, 287–302

Kinloch A J, Shaw S J, Todd D A, Hunston D L 1983 Deformation and fracture behavior of a rubber toughened epoxy. *Polymer* **24**, 1341–63

Majumdar B S, Matikas T E, Miracle D B 1998 Experiments and analysis of single and multiple fiber fragmentation in SiC/Ti-6Al-4V MMCs. *Composites B* **29**, 131–45

Majumdar B S 1998 Fiber-matrix interface. In: Mall S, Nicholas T (eds.) *Titanium Matrix Composites*. Technomic, Lancaster, PA, pp. 113–68

McKenna G B, Mandell J F, McGarry F J 1974 Interlaminar strength and toughness of fiberglass laminates. In: *Proc. SPI Annu. Tech. Conf. RPD*. Society of Plastics Industries, New York

Paris P C, Sih G C 1964 Fracture toughness testing and its applications. ASTM STP 381. American Society for Testing and Materials, Philadelphia, pp. 30–81

Pearson R A, Yee A F 1989 Toughening mechanisms in elastomer modified epoxies—III: effect of cross-link density. *J. Mater. Sci.* **24**, 2571–80

Phillips D C 1972 The fracture energy of carbon fiber reinforced glass. *J. Mater. Sci.* **7**, 1175–91

Reedy E D 1980 Analysis of center notched monolayers with application to boron/aluminum composites. *J. Mech. Phys. Sol.* **28**, 265–86

Tada H, Paris P C, Irwin G 1973 *The Stress Analysis of Cracks Handbook*. Del Research Corporation, Hellertown, PA

Thouless M D, Evans A G 1988 Effects of pull-out on the mechanical properties of ceramic-matrix composites. *Acta Metall.* **36** (3), 517–22

Timoshenko S, Goodier J N 1951 *Theory of Elasticity*, 2nd Edn. McGraw Hill, New York

Tirosh J 1973 The effect of plasticity and crack blunting on the stress distribution in orthotropic composite materials. *J. Appl. Mech., Trans. ASME* **40**, 785–90

Warrier S G, Majumdar B S , Miracle D B 1997 Interface effects on crack deflection and fiber bridging behavior during fatigue crack growth in titanium matrix composites. *Acta Metall.* **45** (12), 4969–80

Wu W L, Hu J, Hunston D L 1990 Structural heterogeneity in epoxies. *Polym. Eng. Sci.* **30** (14), 835–40

Zarzour J F, Paul A J 1992 On plastic zones and fracture strengths in some metal matrix composites. *J. Mat. Eng. Perform.* **1** (5), 659–68

B. S. Majumdar[a] and D. Hunston[b]
[a] *New Mexico Tech, Socorro New Mexico, USA*
[b] *National Institute of Standards and Technology, Gaithersburg, Maryland, USA*

Conventional and Super Abrasive Materials

Abrasives are used extensively for high material removal rate applications as well as for the production of the desired shape, size, finish, and accuracy for a wide range of workmaterials including metals (soft versus hard, ferrous alloys versus nonferrous alloys), ceramics, glasses, wood, concrete, stone, and composites. They differ from conventional cutting tools in that the geometry of the abrasives is random to begin with and changes continuously and in an unpredictable manner as the process continues. In contrast, conventional cutting tools, be it single or multiple point, possess geometrically defined edges whose geometry is more or less maintained the same throughout the process.

1. General

The term *abrasive* is derived from the word *abrasion* meaning one that abrades or scratches another softer material. Scratch hardness is, in fact, one of the oldest forms of hardness measurement. It depends on the ability of one material (in this case a mineral) to scratch another material or mineral. Moh in 1922 introduced an arbitrary, not a linear, semi-quantitative scale for scratch hardness, known popularly as the Mohs scale of hardness. Moh selected 10 minerals as standards, beginning with talc, an extremely soft material (with an arbitrary scratch hardness of one) and ending up with diamond, an extremely hard (the hardest material known) with an arbitrary scratch hardness of 10 (see Table 1).

This way, a mineralogist can thus handily carry the 10 standard minerals to the field to determine the relative hardness of any unknown mineral in the field without the need for elaborate instrumentation. Accordingly, a mineral with a higher Mohs number, for example, quartz (Mohs hardness of seven) can scratch all the minerals directly below it in the hardness scale but not the ones above it. Unfortunately, most hard materials of importance for use as abrasives and cutting tools are in the range of 9–10. Consequently, it is extremely difficult to distinguish them in the Mohs scale. Hence, alternate hardness scales, known as Knoop or Vickers hardness scale, were developed for this purpose based on the principle of indentation.

Grinding operations can be broadly classified into two categories, namely, *form and finish grinding*

Table 1: Mohs hardness scale.

Material mineral	Mohs hardness or scratch hardness scale
Talc	1
Gypsum	2
Calcite	3
Fluorite	4
Apatite Feldspar	5
Orthoclase	6
Quartz	7
Topaz	8
Corundum	9
Diamond	10

(FFG) and *stock removal grinding* (SRG). In FFG operations, such as surface grinding, internal grinding, cylindrical grinding, and centerless grinding, a medium-to-fine abrasive is used while in SRG operations, such as snagging or conditioning of castings in the foundry and abrasive cut-off operations, a coarse abrasive grain is used. The emphasis in FFG is the generation of the required form, finish, accuracy, and surface integrity with less emphasis on the material removal rate, while in SRG, it is just the opposite, namely, the emphasis is on the stock removal rate. Thus, the abrasives used in FFG are medium-to-fine in grain size and very friable while the abrasives used in SRG are generally coarse-to-medium and very strong. Also, abrasives used in the FFG operations are dressed periodically while that in the SRG operations are rarely dressed. Dressing involves the removal of metal chips occupying in the void spaces between the grains in the grinding wheel, removal of the worn flats on the abrasive grains by microfracture and the generation of sharp cleavage fracture facets on the abrasives. This can be accomplished by using a single point diamond dresser, or a multipoint grinding wheel, or an abrasive stick that removes the loading of the wheel and worn flats on the abrasives leaving the wheel with sharp fracture facets and adequate space for the temporary storage of the chips during grinding so that the grinding process can proceed efficiently. The self sharpening action of an abrasive during grinding as well as the generation of sharp fracture facets during dressing depend significantly on the friability of an abrasive. The more friable the abrasive is, the easier it is to generate the self-sharpening action during grinding as well as during dressing. This is the reason why single crystal, refractory ceramics are used as abrasives instead of ceramics or carbides held together with a metal binder, such as cemented tungsten carbide or cermets used in the cutting tools. Abrasives used in FFG are dressed periodically while those in SRG are never dressed. Consequently, the abrasives used in FFG are medium to fine grain and very friable while in SRG, they are generally coarse grain and very strong. FFG is characterized by high wheel speeds (1000–2000 $\mathrm{m\,min^{-1}}$ in conventional grinding and up to 8000 sfpm in high-speed grinding), fine chip thickness values (1–150 μm), and table speeds in the range of 15–18 $\mathrm{m\,min^{-1}}$.

Chemistry and microstructure of an abrasive (including the defect structure) play an important role in its performance. Various properties of interest of abrasives include the hardness, melting temperature, strength and toughness (or its inverse, namely, friability), and thermal conductivity. The efficiency of a material removal process using an abrasive, be it grinding, polishing, honing, etc., depends on the sharp fracture facets on the abrasives that act as cutting points in the removal of material. The ability of these grains to self-sharpen as a result of cleavage or microfracture during the process is the key to the success of an abrasive. Thus defect structure in the abrasive, be it impurities, or voids formed due to the gases evolved during the synthesis of the material, plays an important role in the performance and efficiency of the process. At the same time, this minimizes the cost of production of the abrasives as it is not required to produce pure or defect-free material in contrast to the cutting tool materials where the requirements demand that they be produced with minimal defects and impurities thereby increasing their strength, toughness, hardness, and wear resistance.

The character of an abrasive depends on the method of comminution or the size reduction process used. In the case of conventional abrasives, generally, huge lumps of material produced by fusion are reduced to different sizes using different types of jaw crushers, impact crushers, and ball mills. They are subsequently sieved to separate them according to the grain size required. While crushing to the required grain size is the only practical method presently used to reduce grain size of most abrasives, it frequently introduces micro- and even macrocracks which can be quite objectionable particularly for large and medium sized grains. When the grain size under consideration is very small, defects due to microcracks disappear and such grains are found to be very strong and yet exhibit sharp cleavage facets which are ideal for material removal operations by abrasives. In some cases, flaws introduced in crushing are removed by a crack-healing thermal treatment. There are, however, other methods by which abrasives can be produced to the desired size without the need for crushing or subjecting it to minimal size reduction processes.

Grain size of an abrasive plays an important role in the material removal and surface generation process. The grain size, G, is related to the mean diameter of the abrasive grain, d (Fig. 1). It is given by an approximate relationship of the form, $G\ d = 0.6$, where d is in inches. Thus, a grain size of 100 would have a mean diameter of ~ 0.15 mm. Sieves with different size openings are used for sorting out abrasives by the grain size. Thus an abrasive of 100 grain size would pass through a sieve with 100 openings per 6.25 $\mathrm{cm^2}$. In the specification of grinding wheels, a grain size in the range of 8–24 is considered coarse, 30–60 coarse, 70–180 medium, and 220–600 fine to very fine. A similar classification exists for the use of abrasives in loose form as in polishing and lapping as well as in coated form. The abrasive grains used in coated abrasives generally have a long aspect ratio or needle-like shape to enable stronger bonding of the abrasive to the backing material and to obtain longer life. These abrasives are oriented along their length disection.

Abrasives are used in both conventional machining processes, such as grinding, polishing, lapping, honing as well as nonconventional processes, such

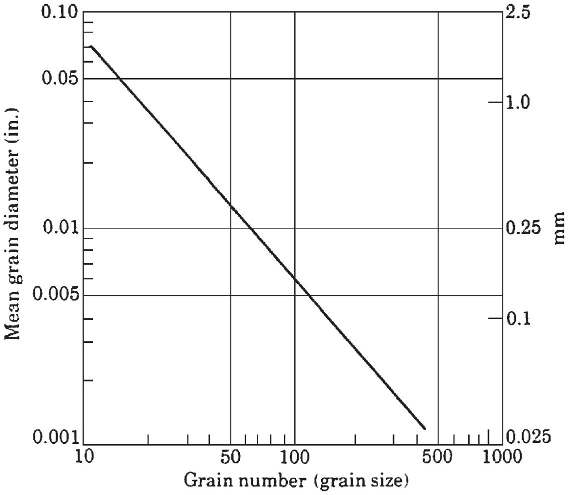


Figure 1
Relationship between the grain size and mean grain diameter of abrasives.

as ultrasonic machining, abrasive jet machining, magnetic field assisted polishing. Abrasives are used in three different forms (i) bonded abrasive form, such as in grinding wheels, honing sticks; (ii) coated abrasive form, such as in coated abrasive paper or cloth (popularly known as *sand paper* although other abrasives, such as alumina, granet, and SiC are often used); and (iii) loose abrasive form, such as in polishing, and lapping. The focus of this article is on abrasives both conventional and superabrasives.

The performance of a grinding wheel in grinding a given workmaterial depends on the type of abrasive used, its grain size, the wheel hardness (soft to hard, depending on the amount of the binder used and represented by an alphabet A to Z), wheel structure (dense to open, represented by a numerical value 1–15), and the type of bond. Figure 2(a) and (b) are ANSI specifications of grinding wheels used with conventional abrasives and superabrasives, respectively. In the case of conventional abrasive, the specification gives the type of abrasive used (Al_2O_3 or SiC), the grain size, the grade (which measures the relative hardness or the strength with which the abrasive grains are held by the bond), the structure (the openness or the spacing between the grains), and the type of bond used. Both the hardness and the structure depend on the amount of binder content used.

A typical vitrified bonded conventional abrasive grinding wheel would consists of ~50% abrasive, ~10% bond, and ~40% porosity. Voids (or porosity in the wheel) are intentionally created for the temporary storage of chips by incorporating a volatile filler material in the initial powder mix. For example, sawdust or a material which sublimes on firing, such as hexachlorethane ($C_2H_2Cl_6$) is introduced in the mix before pressing the wheel. A typical vitrified bond material may consist of ~63% SiO_2, 12% Al_2O_3, 1.2% CaO, 6.3% Na_2O, 7.5% K_2O, and 10% B_2O_3 (by weight). The abrasive, bond material, and saw dust are mixed with temporary binders to hold it in the green state and the grinding wheels are pressed to shape in steel molds. They are then removed from the molds and fired at ~900–1300 °C, held at this temperature for ~8 hours, and cooled down to room temperature. When these materials volatize, they leave behind voids in the grinding wheel. The final product is inspected by such tests as measurement of the ratio of mass to volume, modulus of elasticity, resistance to penetration by rotating a chisel point or by a blast of known volume of sand or other abrasive under a known pressure.

Phenol formaldehyde bonded alumina-based or alumina–zirconia-based wheels make up for the largest part of the organic bonded or resinoid bonded grinding wheels for SRG. For high-speed (~5000–8000 m min^{-1}) applications, they are usually reinforced with fiberglass. Resin bonded wheels can be phenol formaldehyde, epoxy, polyurethane, polyester, or shellac. The bond usually includes suitable filler materials, such as cryolite, iron sulfide, CaF_2, ZnF_2, and ammonium chloride. In the case of resinoid bonded superabrasive (diamond or cBN) wheels, a rim of resin bonded abrasive is used on top of a nonferrous disk. Hence, in the specification, its thickness is given. Also, the concentration of the diamond (or cBN) in the resin is also given. This can vary from 25% to 50% depending on the application and cost. A 100% concentration corresponds to 72 carats of diamond (or cBN) in one cubic inch of the active wheel medium (i.e., rim of the resin bonded abrasive). Also, five carats of the superabrasive is equivalent to one gram.

Of the various types of binders used in the manufacture of grinding wheels, glass or vitrified bond (V), resin bond (B), and metal bond (M) are the most common. Vitrified bond is used with conventional abrasives (alumina and SiC) in FFG and resin bond in SRG. With superabrasives (diamond and cBN) a resin bond is used extensively for grinding wheels while a metal bond is used for saw blades. Electroplated wheels contain a monolayer of the superabrasive mounted on the periphery of a metal disk. As the disks have already been machined concentric, there is no need for dressing or truing the wheel which is rather critical in the performance of superabrasive wheels. In addition, the cost of the wheel is only a fraction of a resin bonded or a metal bonded wheel. Vitrified bond is used for selective applications of both diamond and cBN wheels. Soft grinding wheels with open structure (i.e., containing less amount of the binder) are used for machining hard workmaterials while hard wheels with dense structure (containing more amount of the binder) are used for machining soft workmaterials.

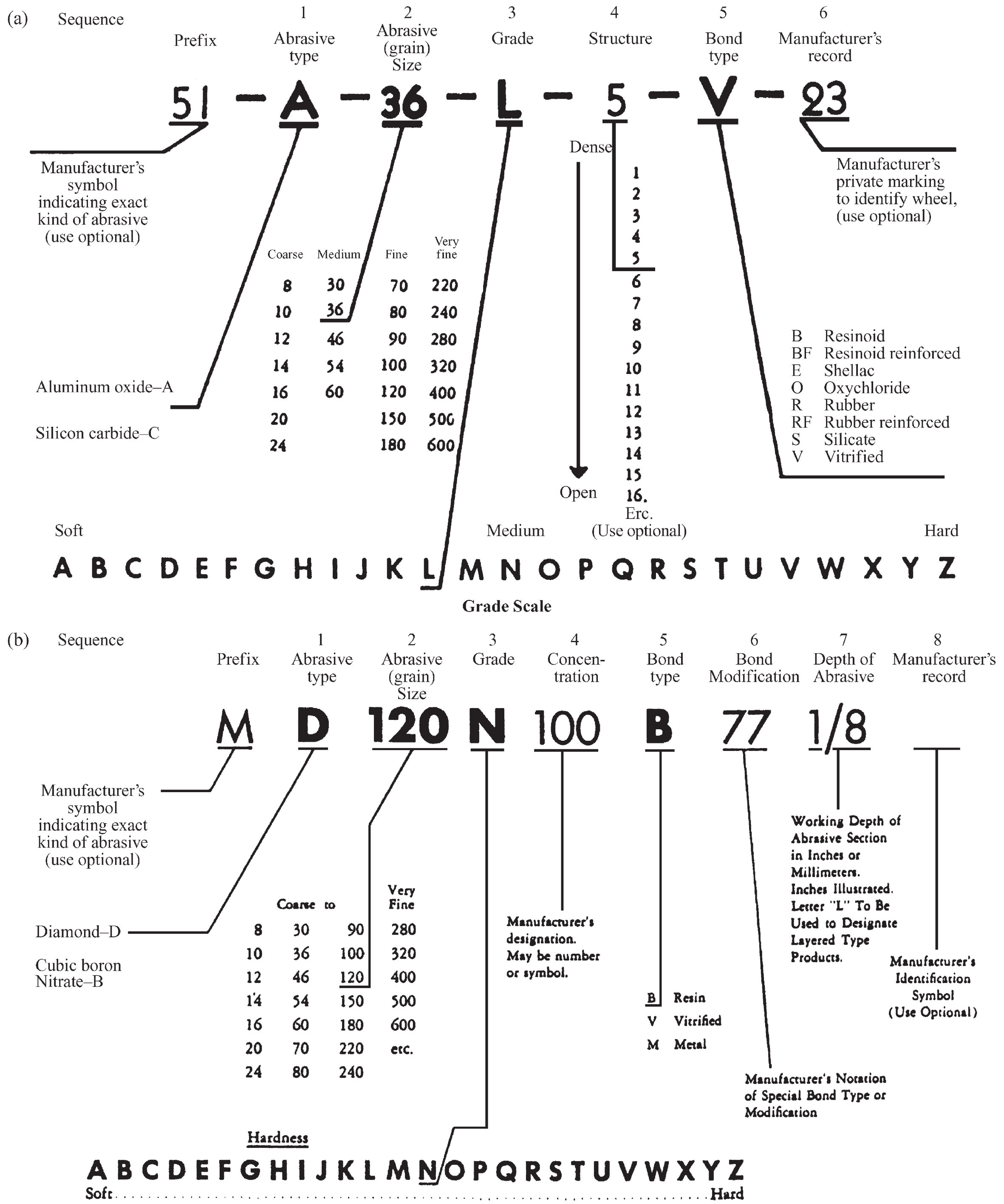

(a)
Sequence
Prefix
1 Abrasive type
2 Abrasive (grain) Size
3 Grade
4 Structure
5 Bond type
6 Manufacturer's record
51 – A – 36 – L – 5 – V – 23
Manufacturer's symbol indicating exact kind of abrasive (use optional)
Aluminum oxide–A
Silicon carbide–C
Coarse
Medium
Fine
Very fine
8 10 12 14 16 20 24
30 36 46 54 60
70 80 90 100 120 150 180
220 240 280 320 400 500 600
Dense
Open
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16.
Etc.
(Use optional)
B Resinoid
BF Resinoid reinforced
E Shellac
O Oxychloride
R Rubber
RF Rubber reinforced
S Silicate
V Vitrified
Manufacturer's private marking to identify wheel, (use optional)
Soft
Medium
Hard
A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
Grade Scale
(b)
Sequence
Prefix
1 Abrasive type
2 Abrasive (grain) Size
3 Grade
4 Concen-tration
5 Bond type
6 Bond Modification
7 Depth of Abrasive
8 Manufacturer's record
M D 120 N 100 B 77 1/8
Manufacturer's symbol indicating exact kind of abrasive (use optional)
Diamond–D
Cubic boron Nitrate–B
Coarse to
Very Fine
8 10 12 14 16 20 24
30 36 46 54 60 70 80
90 100 120 150 180 220 240
280 320 400 500 600 etc.
Manufacturer's designation. May be number or symbol.
B Resin
V Vitrified
M Metal
Manufacturer's Notation of Special Bond Type or Modification
Working Depth of Abrasive Section in Inches or Millimeters. Inches Illustrated. Letter "L" To Be Used to Designate Layered Type Products.
Manufacturer's Identification Symbol (Use Optional)
Hardness
A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
Soft .... Hard

Figure 2
ANSI specifications of grinding wheels using (a) conventional abrasives (Al_2O_3 and SiC) and (b) superabrasives (cBN and diamond) in various bonds.

1.1 Coated Abrasives

Coated abrasives are available in a variety of forms including sheet, disk, roll, and belt and can be used for a range of applications from very light to heavy duty. A coated abrasive consists of an abrasive held in a backing material by an appropriate bonding medium (Fig. 3). The backing material serves as the base upon which a coating of an adhesive known as the *make* coat is uniformly applied to anchor a single layer of an abrasive. A typical make coat comprises of a solution containing ~48% phenol formaldehyde resin and ~52% calcium carbonate filler. Usually, abrasive grains with a long aspect ratio are specially prepared and used for this application (Fig. 4). Individual abrasive particles are applied uniformly to the "make" coat and oriented by an electrostatic field to maximize the probability that an abrasive will be positioned with its major or long axis perpendicular to the backing material. After the make coat hardens (generally precured for ~30 minutes at 107 °C), a second coating of adhesive known as the "size" coat (of a similar composition as the make coat) is applied for anchoring the abrasives strongly to the backing material and cured for ~10 hours at 107 °C. A wide range of abrasives both natural and synthetic is used for coated abrasive applications. They include conventional abrasives such as the various alloys of alumina, sintered sol-gel alumina, SiC, emery, garnet, flint as well as superabrasives, diamond, and cBN. They are available in all grit sizes (36–600) for a wide range of applications. The coated abrasive can be either open-coat abrasive (~50–70% of its backing surface is covered with the abrasive) for free cutting action or closed-coat (>70%) for obtaining high removal rates.

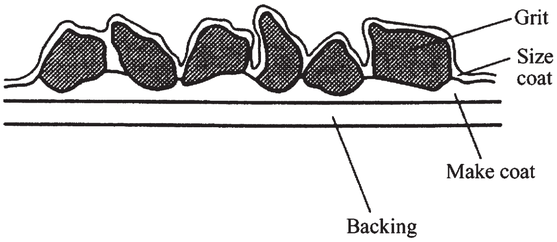


Figure 3
Schematic of a coated abrasive product showing abrasive grits anchored to the backing material (paper, cloth) with a "make" coat initially followed by a "size" coat.

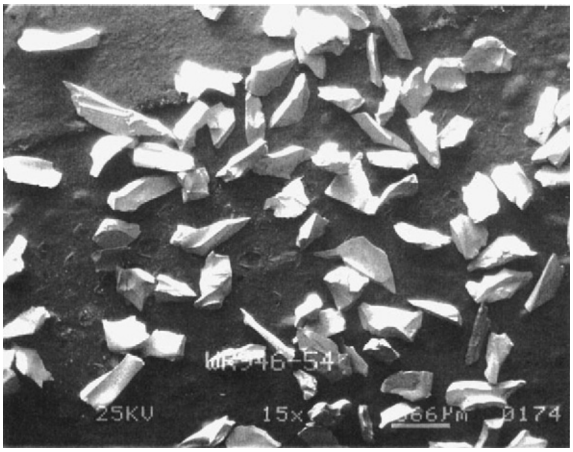


Figure 4
SEM micrograph of a 46 grit sol-gel alumina abrasives used in coated abrasive applications showing abrasives with a long aspect ratio produced by a special crushing operation.

1.2 Honing

Honing is a finishing process used for the generation of functional characteristics specified for a surface, including geometric and dimensional accuracy as well as surface requirements, such as roughness, lay pattern, and surface integrity. It is a low speed abrading operation using bonded abrasive sticks inserted into a tool holder. The sticks are spring loaded to expand away from the tool to grind irregularities in the surface of the part. The process involves both rotational and reciprocating action and generates a cross-hatched pattern. Honing can correct out-of-roundness, taper, bell shaped, and barrel shaped bores. Honing sticks are available in all grits sizes (36–600) of SiC, alumina, and superabrasives (diamond and cBN) and in all types of bonds including vitrified, resin, and metallic depending on the application.

1.3 Lapping

Lapping is a finishing operation used for the production of flat surfaces. The abrasive is applied on to a cast iron lap of a lapping machine in the form of an abrasive slurry in a water-soluble glycol solution. Alternately, a coated abrasive product can be used on a backing plate. A continuous feed system ensures fresh supply of the abrasive. To provide uniform lapping, the workpiece is loaded to provide uniform pressure and is rotated during the lapping operation.

Although abrasives, in general, are hard, brittle materials they can be softer than the workmaterial in which case the material removal or finish generated does not involve mechanical abrasive action but chemical- or chemo-mechanical action. Abrasives can be either natural minerals or synthetic materials.

Both are extensively used depending on the type of application although synthetic materials are rapidly replacing, in many cases, because of their ready availability, reliability, consistency, controlled properties, and cost benefits. The most common natural minerals include quartz, or silica sand, zircon sand, corundum, emery, garnet, flint, and diamond. Corundum and emery are natural forms of aluminum oxide containing varying amounts and nature of impurities. Many of the natural abrasives, because of varying amounts of impurities, tend to perform nonuniformly and, hence, are used only for those applications where this is not a serious limitation but the low cost and easy availability of the abrasive are beneficial factors. Synthetic materials include various types of aluminum oxide and its alloys, silicon carbide, boron carbide, zirconium oxide as well as alumina–zirconia, cubic boron nitride (cBN), and diamond. Chemo-mechanical abrasives include, cerium oxide (CeO_2), magnesium oxide (MgO), silica (SiO_2), chromium oxide (Cr_2O_3), alumina (Al_2O_3), titania (TiO_2), zirconia (ZrO_2), iron oxide or rouge (Fe_3O_4), and yttrium oxide (Y_2O_3). Their performance depends on the interactions between the abrasive–workmaterial–environment as well as on the conditions of polishing, such as the temperature and the pressure at the contacting zone.

In the following, various types of natural and synthetic abrasives as well as conventional abrasives and superabrasives will be briefly described along with some of their applications in different manufacturing operations.

2. *Natural Abrasives*

Natural abrasives possess varying amounts of impurities and defect structure. Hence, their performance generally is somewhat nonuniform and unpredictable. But, because of their limited volume usage or intrinsic value, their synthesis is not justified for they perform satisfactorily and economically in certain operations.

Quartz sand (SiO_2) or flint is used both in the loose form and in coated abrasive form. The former was used for the grinding of plate glass (prior to the development of float glass manufacturing technology) and is still used for sand blasting mostly in foundries for cleaning the castings and in the preparation of metal surfaces prior to painting. Flint is used as a coated abrasive paper for the sanding of wood. Quartz is hexagonal in structure and has a distinct rhombohedral cleavage and a conchoidal fracture characteristics. Its Knoop hardness is $\sim 820\,kg\,mm^{-2}$.

Garnet is an orthosilicate of iron (or magnesium) aluminum silicate. Two common minerals of garnet are (i) almandite with a chemical formula of $Fe_3Al_2(SiO_4)_3$ and (ii) pyrope garnet with a chemical formula of $Mg_3Al_2(SiO_4)_3$. Hardness of garnet is $\sim$7.5–8.5 on the Mohs scale or $1360\,kg\,mm^{-2}$ on the Knoop scale with a decomposition temperature of $\sim$1100 °C. The color of garnet ranges from dark red to brownish red. Though superior to quartz in performance, it is higher in cost and used for similar applications. In the coated abrasive form it is used extensively especially in fine grinding of softwood. It is also used in the finishing of glass, stone, leather, and varnished and painted surfaces. In loose form, it is used in abrasive jet processes, such as sand blasting. The intrinsic value of garnet is determined more by its friability (which is its ability to cleave easily along certain weak crystallographic planes and directions to provide sharp fracture facets) than its hardness.

Feldspars are a class of aluminum silicate minerals containing K, Na, or Ca or a combination. They are a major component of igneous and metamorphic rocks. The rocks are mined, crushed, screened and ground to size. The hardness of the feldspar is $\sim$6 on the Mohs scale. They break with a semi-conchoidal fracture that make them useful as a mild abrasive. A major use of fledspar is in scouring powders. They are not as hard as quartz, so they are safer to use on glass and enamel surfaces.

Pumice is a light colored volcanic glass produced by explosive volcanic eruptions. It is used as an abrasive in the form of fine powder for polishing metals and in finishing painted surfaces.

Corundum is a natural form of α-Al_2O_3 available in various states of purity and crystallinity formed by contact metamorphism. Its hardness is $\sim$9 on the Mohs scale and its color varies considerably from blue, green, gray, brown, and black to colorless. Transparent blue sapphire and the red ruby are valuable gemstones of natural corundum. The natural forms of alumina are distinguished as corundum, emery, and spinel (γ-alumina with a spinel structure). Natural corundum crystallizes in the rhombohedral form and has a basal cleavage not found in the synthetic counterpart, fused alumina. As an abrasive, corundum is used primarily in the grinding and polishing of glass lenses. However, much of it use has been replaced by its counterpart, synthetic alumina.

Emery is an impure form of corundum, a natural aluminum oxide in which the impurity is mostly magnetite or spinel depending on the location of the available deposits. The hardness ranges from 7.5–8.5 on the Mohs scale. Processing involves washing, crushing, and sizing, followed by heat treatment. Coarse emery is extensively used in nonskid concrete floors and pavements and to produce millstones for rice milling. The medium-to-finer grades are used to make emery paper and cloth, and the very fine goes in the production of polishing pastes.

Natural diamond is a class by itself amongst all abrasives. Its unique properties include the highest hardness, excellent wear resistance, good thermal conductivity, easy cleavage along certain crystallographic planes and directions, and extremely low friction. Its applications extend from the finishing of diamond itself to the cutting, grinding, and finishing of a range of hard materials, such as stone, concrete, glass, hardened alloy steels, cemented carbides, and ceramics. It is an indispensable and strategically

Table 2
Chemical analysis, percent weight.

Abrasive material	Knoop hardness (kg mm^2)	Fe	Fe_2O_3	Si	SiO_2	Ti	TiO_2	Na	SO_3	C	S	Loss on ignition	ZrO_2	MnO	CaO	MgO
White aluminum oxide	2122	0.03		0.02		2.01		0.25	<0.01	0.07		0.07				
Monocrystalline aluminum oxide	2276		0.10		0.11		0.61				0.17		0.11	<0.005	0.05	0.02
Chromium modified aluminum oxide	2258															
Vanadium modified aluminum oxide	2350															
Regular aluminum oxide	2042		0.21		0.73		2.65									
40% zirconia alloyed aluminum oxide	1462		0.25		1.84		1.44						42.05			
Sintered bauxite	1372		5.72		3.21		3.31						0.06	0.01	0.18	0.26
Black silicon carbide	2679			1.29	1.96						0.83	0.66			0.01	0.03
Green silicon carbide	2837			0.63	1.18						0.65	0.4			0.02	0.03

important material whose role cannot be replaced except by itself in the synthetic or man-made form. Until the successful synthesis of diamond at high pressure–high temperature (HP-HT) in the mid1950s and subsequent synthesis of cubic form of boron nitride (cBN) (which incidentally does not exist in nature) natural diamond occupied a strategic role in the manufacturing field. Similar to all other natural mineral abrasives, the performance of natural diamond depends on the impurity content as well as the defect structure in the material. Synthetic diamond (as well as cBN) with more control on the impurity content as well as crystal morphology (similar to the other conventional abrasives), is leading its natural counterpart in its application and use in manufacturing. It is amazing that the synthetic materials (diamond and cBN) are not only more controlled in terms of chemistry, crystallinity, and defect structure but are also available in unlimited quantities economically.

Silicon carbide has the distinction of being the first commercial synthetic abrasive while cBN has the unique distinction of a synthetic ultrahard material (or superabrasive) that does not even have a natural counterpart and whose synthesis is based on the similarities in the crystal structures between the hexagonal forms of carbon (graphite) and boron nitride prior to the synthesis and their cubic counterparts (diamond and cBN) after the synthesis.

Various types of synthetic abrasives are used in the manufacturing industry. They can be classified as alumina-based (white alumina, monocrystalline alumina, regular alumina, sintered bauxite, sol-gel alumina, Al_2O_3–ZrO_2), SiC, B_4C, cBN, and diamond. Alumina based abrasives are used for the finishing of steels and SiC abrasive is used for the finishing of nonferrous alloys as well as cast iron. Diamond is used for all hard materials, such as stone, concrete, cemented carbide, ceramics, glass, or silicon but is not recommended for low carbon steels for it reacts chemically with Fe. cBN is used for the finishing of hardened steels (R_c 45–65) and nickel-base superalloys.

3. *Synthetic Abrasives—Alumina-based*

Aluminum oxide (Al_2O_3) has a hexagonal crystal structure ($a = 4.758$ Å, $c = 12.99$ Å and $c/a = 2.73$), a melting temperature in the range of 2015–2050 °C and a hardness $\sim 2100\,\text{kg}\,\text{mm}^{-2}$ depending on the purity. Various types of aluminum oxide abrasive are used in industry including white aluminum oxide, monocrystalline aluminum oxide, regular aluminum oxide, modified (chromium oxide or vanadium oxide) aluminum oxide, sintered bauxite or sintered alumina, alumina–zirconia, and sintered sol gel alumina. Table 2 gives the hardness and chemical composition of various types of alumina-based abrasives.

White aluminum oxide is manufactured either by the fusion of alumina or using the Bayer processed (dissolving the bauxite in caustic soda and then re-precipitating it to remove all impurities) alumina. The first method uses a furnace of the Higgins-type made of a water cooled, truncated cone shaped steel vessel with carbon electrodes (Fig. 5). Graphite or coke is used at the beginning to serve as a conducting path. Initially ~20–25% of charge containing a mixture of 80% bauxite, 15% iron filings, and 5% coke is added to the furnace. The purpose of the coke is to reduce the silica to silicon and the iron oxides to iron. Iron then combines with silicon to form ferrosilicon which settles at the bottom due to higher density and can be easily separated by magnetic means. Once fusion starts, the initial charge acts as the conducting path and the rest of the charge is added over a period of 15–30 hours. depending on the size of the furnace.

In the second method, bauxite is treated with sodium hydroxide (Na_2O) at a temperature of 120–140 °C to form sodium aluminate, which is filtered from insoluble impurities. The precipitate is then washed to remove soda and then calcined at 1100 °C and fused in Higgins type of furnace. A typical Higgins furnace is ~1.8 m in diameter and 1.8 m in height. It operates on a 3-phase, 60 Hz current at an applied voltage of ~100–150 V and a power of ~750 kW. The furnace is under power for ~18 hours. The resulting ingot weighs ~5442 kg. When fusion is complete, the electrodes are removed and the molten material is allowed to cool for 3–4 hours in the furnace and about a week outside the furnace. The slow cooling produces coarse crystallites. If finer crystallites are desired, it is poured into a thin, water cooled, steel container for rapid cooling. The ingots are crushed and sieved to size to obtain the abrasive of various grain sizes.

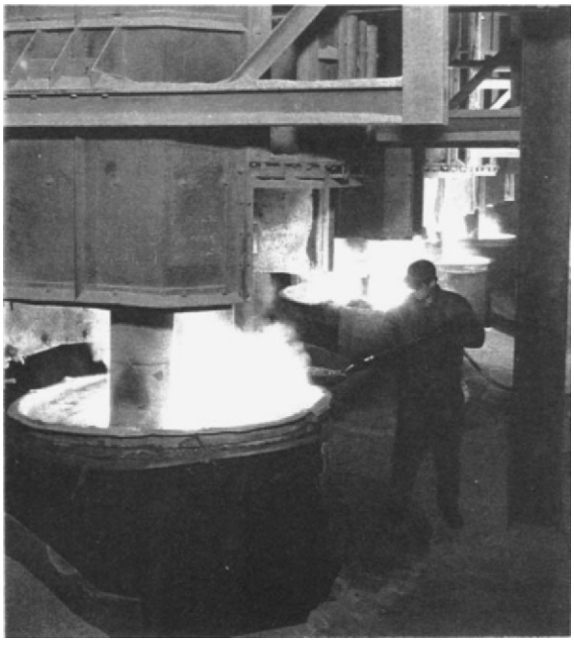


Figure 5
Photograph of a Higgins-type furnace used for the fusion of Al_2O_3 from bauxite.

As can be seen from Table 2 (Komanduri and Shaw 1974b), this material is very pure except for the small amounts of residual sodium hydroxide (Na_2O) (0.25–0.5%) that volatilizes during melting giving rise to a porous structure. Figure 6(a) is an SEM micrograph of a coarse grain (12 grain size) white aluminum oxide showing considerable number of gas holes due to the formation of Na_2O. However, when the abrasive it reduced by comminution to a finer grain size, the crystal breaks through these imperfections resulting in sharp fracture facets that are characteristic of alumina abrasive (Fig. 6(b)).

Monocrystalline aluminum oxide abrasive is produced by the fusion of a mixture of bauxite, coke, iron sulfide, iron filings, and alkaline metal oxides. The resulting ingot consists of hydrolyzable alumina in a decomposable sulfide matrix. In this process, the fused ingot is cooled slowly to crystallize as numerous single crystals. A light crushing operation followed by water treatment that disintegrates the glassy matrix and leaves the abrasive in loose, single crystal form. The material is free from Na_2O and hence is free from gas holes.

As shown in Table 2, this abrasive contains small amounts of (up to 0.1%) Fe_2O_3, SiO_2, and ZrO_2 and up to 0.6% TiO_2. These oxides alloy with the base alumina and improve the properties of this material. This abrasive is characteristically very light pink in color. Figure 7(a) and (b) show SEM micrographs of an 80 grain size monocrystalline aluminum oxide abrasive at two magnifications showing a series of sharp stepwise cleavage facets formed during the manufacture. Such a structure can be seen throughout the size range.

Regular aluminum oxide, also known as high titania fused alumina, comprises a coarse crystalline product having typically ~2.65% residual TiO_2 as shown in Table 2. It is manufactured in a batch process by the fusion and partial reduction of calcined bauxite (with its associated impurities comprising of titania, silica, and iron oxide) with coke and iron filings in a Higgins furnace. Conditions in the arc furnace during fusion are distinctly reducing. The silica in the bauxite is reduced to silicon and the iron group oxides to metal. Iron then combines with silicon to form ferrosilicon which settles at the bottom due to higher density and is separated by magnetic means. Since, Ti_2O_3 is appreciably soluble in alumina, titanium acts as a modifying element and improves its properties. Although this material has lower hardness and friability than white aluminum oxide, it has higher strength. Hence, it finds applications in abrasive machining ranging from semifinish to rough grinding.

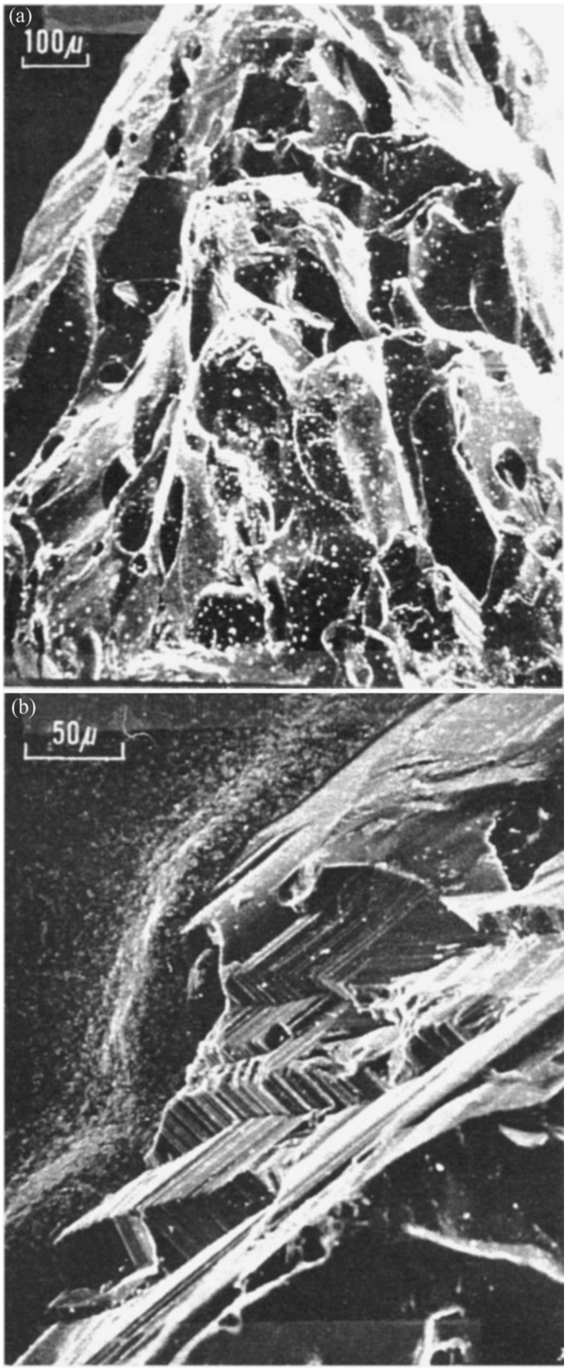


Figure 6
(a) SEM micrograph of a 12 grain size white Al_2O_3 abrasive showing no sharp cleavage facets and considerable porosity due to escape of volatile gases during fusion and (b) SEM micrograph of a 36 grain size white Al_2O_3 abrasive showing sharp cleavage fracture facets (after Komanduri and Shaw 1974a).

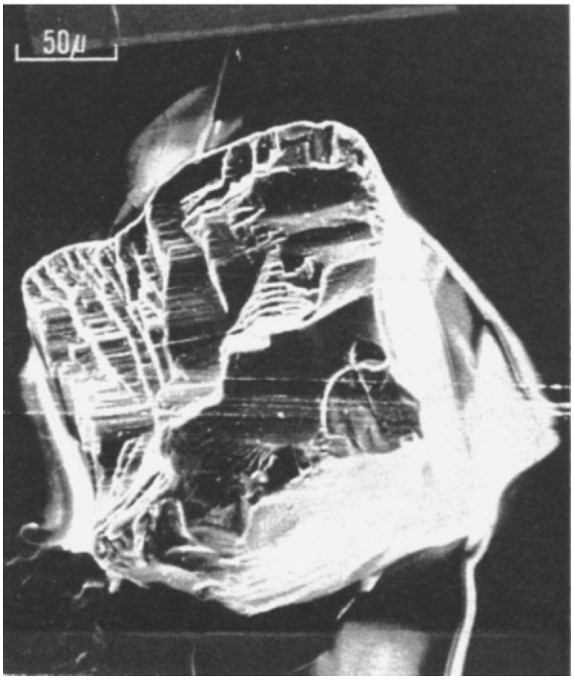


Figure 7
SEM micrograph of a 80 grain size monocrystalline Al_2O_3 abrasive showing a series of sharp stepwise facets formed during growth (Komanduri and Shaw 1974a).

Figure 8 shows an SEM micrograph of a regular aluminum oxide abrasive showing a somewhat rougher surface that can facilitate in the formation of a strong bond in a grinding wheel. The abrasive after fusion and cooling is crushed and sent pan mills for rounding of the edges. The topographical features of this material suggest that it is composed of more than one phase and that one of the phases is more resistant to abrasion than the other. That is, during the pan milling operation, it appears that the softer phase is preferentially removed, thereby generating rougher surface as shown in Fig. 8.

Modified aluminum oxide abrasives are developed because the properties of α-alumina can be modified by significant additions of other alloying oxides in the fusion of calcined Bayer processed alumina. Two such alloying oxides are Cr_2O_3 and V_2O_5. Chromium oxide alloyed with alumina has a pink to ruby color depending on the percentage of the alloy while the vanadium oxide alloyed alumina has a light to dark green color. They are, therefore, identified as ruby and emerald abrasives, respectively. The purpose of tinting is partly aesthetic and partly due to the fact that chromium ions in solid solution in the α-alumina lattice tend to lock slip planes producing a hardening effect. The improved hardness of vanadium oxide alloyed alumina is also attributed to a similar origin.

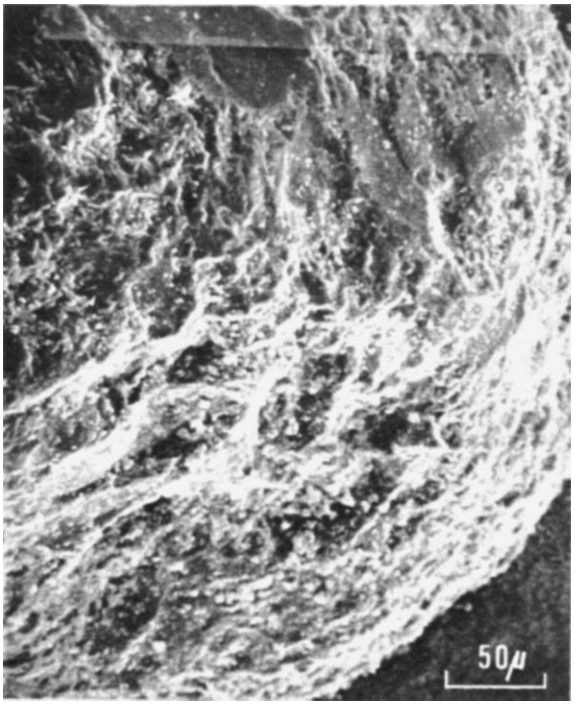


Figure 8
SEM micrograph of a 12 grain size regular Al_2O_3 abrasive showing no sharp cleavage facets (Komanduri and Shaw 1974a).

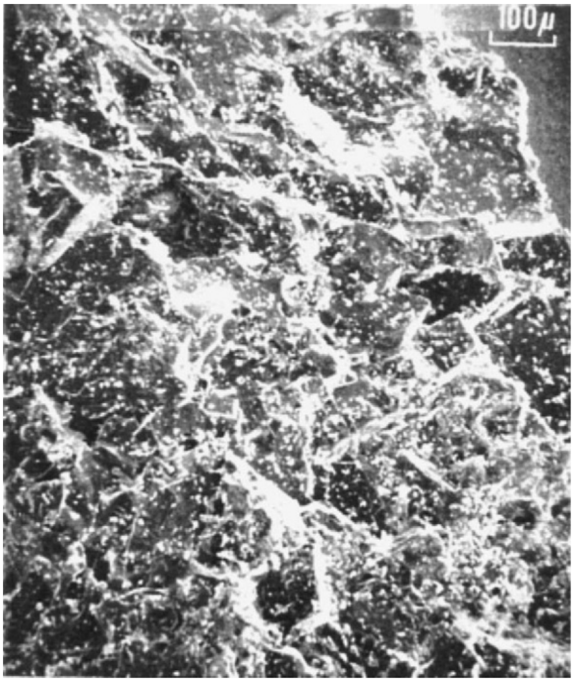


Figure 9
SEM micrograph of a 12 grain size chromia modified Al_2O_3 abrasive showing no sharp cleavage facets but pyramid shaped projections on the surface which can act as cutting points (Komanduri and Shaw 1974a).

Grinding wheels made of this abrasive are slightly more expensive than conventional white or monocrystalline alumina wheels but used in grinding of hardened tools steels and stainless steels with slightly improved performance.

Both Cr_2O_3 and V_2O_5 modified alumina are found to exhibit entirely different topographical features. Figure 9 is an SEM micrograph of a 12 grain size chromia modified alumina. The surface appears to have been formed by intercrystalline fracture and consist of several pyramid shaped sharp edges which can act as cutting edges. Several white spots can also be seen. It appears that as a result of static charge, these grains collect numerous minute specks of the abrasive during comminution. Figure 10(a) is an SEM micrograph of a 12 grain vanadium oxide modified alumina. The grain appears to be a solid mass containing several cavities possibly resulting from entrapped gases formed during the fusion of this alloy. The cavities have the appearance of mini-volcanoes unlike the continuous holes observed in white alumina or SiC. Figure 10(b) is an SEM micrograph at higher magnification showing details of the crater-like appearance of one of the gas pockets. Around the hole, the material has the appearance of an interwoven network although the walls inside the hole show no such texture.

Sintered bauxite or sintered alumina is manufactured by pressing or extruding a finely milled (1–5 μm) paste of calcined bauxite to shape in the form of rods or platelets. They are then chopped to size in the green state and sintered in a rotary kiln at a temperature of ~1450–1500 °C. No sintering aid is necessary for impurities in the calcined bauxite provide this function. Consolidation, diffusion bonding, and grain growth takes place at the sintering temperature. Sintered bauxite has relatively low hardness as a result of high impurity content and microporosity (Table 2). Also, the density of this material is significantly less than theoretical. Nevertheless, it has low friability and high strength. It also does not show cleavage facets; instead the grains are rounded. Hence, sintered bauxite is suitable for conditioning and other heavy duty applications where coarse grain size of this material is used.

Figure 11(a) is an SEM micrograph of a sintered bauxite grain. It contains neither sharp edges nor microvoids as found in the very friable alumina abrasives (e.g., white or monocrystalline alumina). Figures 11(b) and (c) show, at higher magnifications, details of the knobby structure generated by intergranular separation either in the green state or during sintering. The knobby structure closely resembles the

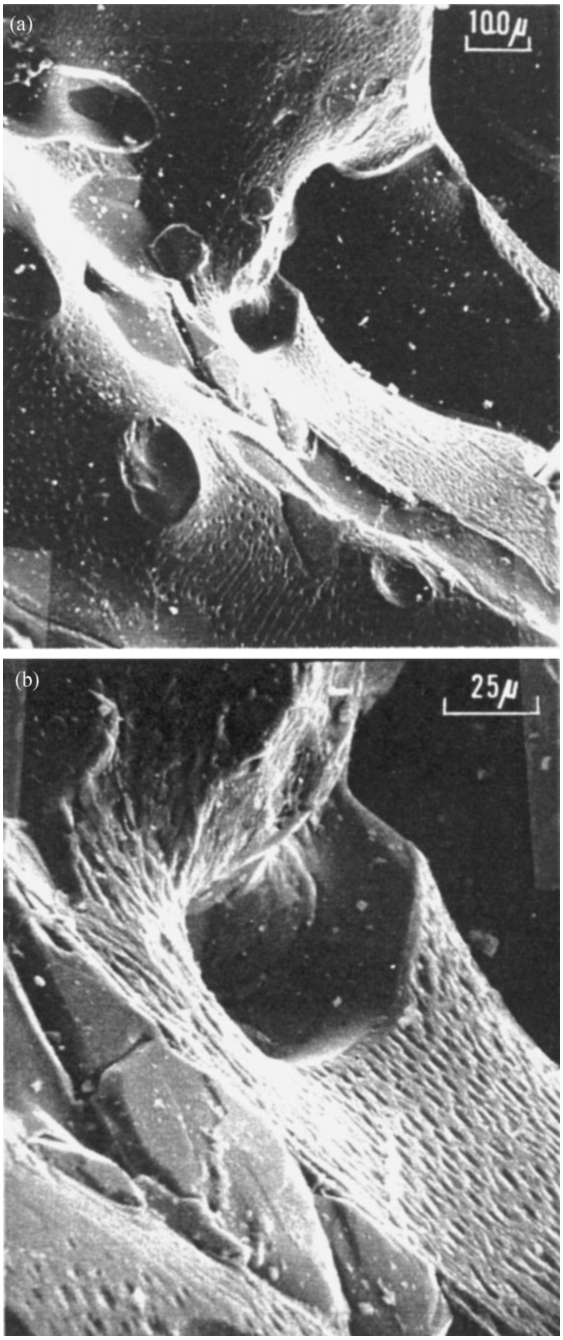


Figure 10
SEM micrographs of (a) a 12 grain size vanadium oxide modified Al_2O_3 abrasive showing several cavities formed due to escape of volatile gases formed during fusion. (b) Details of the inter-like appearance of one of the gas pockets. Around the hole, the material has the appearance of an inter-over network although the walls inside show no texture (Komanduri and Shaw 1974a).

morphology of the finely milled calcined bauxite. The size of the individual crystallites in the sintered product is $\sim 4\,\mu m$. Considerable void space can also be seen, indicative of density less than theoretical. As there is no further comminution, the morphology of sintered bauxite remains the same regardless of the grain size.

Sol-gel microcrystalline alumina abrasive (both unseeded and seeded) is one of the newest developments in the conventional abrasive field for both coated and bonded abrasive applications developed by Norton as well as 3 M company in the USA. It may be noted that the state in which fine particles (nano or submicrometer sized) are dispersed in a liquid is called a *colloid*. When the colloid is sufficiently fluid and stable for a long period of time, it is termed the *sol*. Rigid solids formed by the evaporation of the solvent from a sol is called the *gel*. By manipulating the sol-gel transformation appropriately, it is possible to form a variety of extremely fine-grained shapes rapidly. The essence of this method lies in its ability to produce extremely fine microcrystalline (submicron) abrasive of interest. Such an abrasive provides very high toughness and enables high removal rates in grinding or coated abrasive applications. Various oxide ceramics can be made by the sol-gel method but the current commercial applications are focused on α-alumina and alumina–zirconia abrasives.

Sol-gel microcrystalline alumina abrasive can be produced with or without seeding with fine (submicron) α-alumina particles. The seeds (up to $\sim 5\%$ by weight) provide the nucleation sites for the conversion of the precursor to the α-alumina phase. In general, the seed material should be isostructural with α-alumina with lattice parameters within 15% and be present in the dried gel at the temperature at which the conversion to α-alumina takes place. In the synthesis of seeded sol-gel alumina, submicron sized particles of α-alumina are uniformly dispersed in an aqueous sol or a gel, of an α-alumina precursor, such as boehmite ($Al_2O_3 \cdot H_2O$). To this mixture a peptizing agent, such as HNO_3 in water is added. In order to insure complete peptization of the alumina and homogenization of the gel, the mixture is forced through a kneader/mixer and subsequently extruded. Gelation greatly restricts chemical diffusion and segregation. The mixture is then dried and crushed either before or after sintering to form a roughly graded particulate material. This was then calcined at $>400\,°C$ for ~ 16 hours and then fired for ~ 30 minutes in a rotary kiln to a temperature above the conversion temperature at which α-alumina is formed and sintered essentially to very high density ($>95\%$ theoretical). The sintering temperature is generally in the range of 1300–1400 °C in the case of seeded sol-gel material and about 100 °C more in the case of nonseeded sol-gel. Excessive temperatures or long sintering times can lead to grain growth which should be prevented in order to obtain high toughness of this

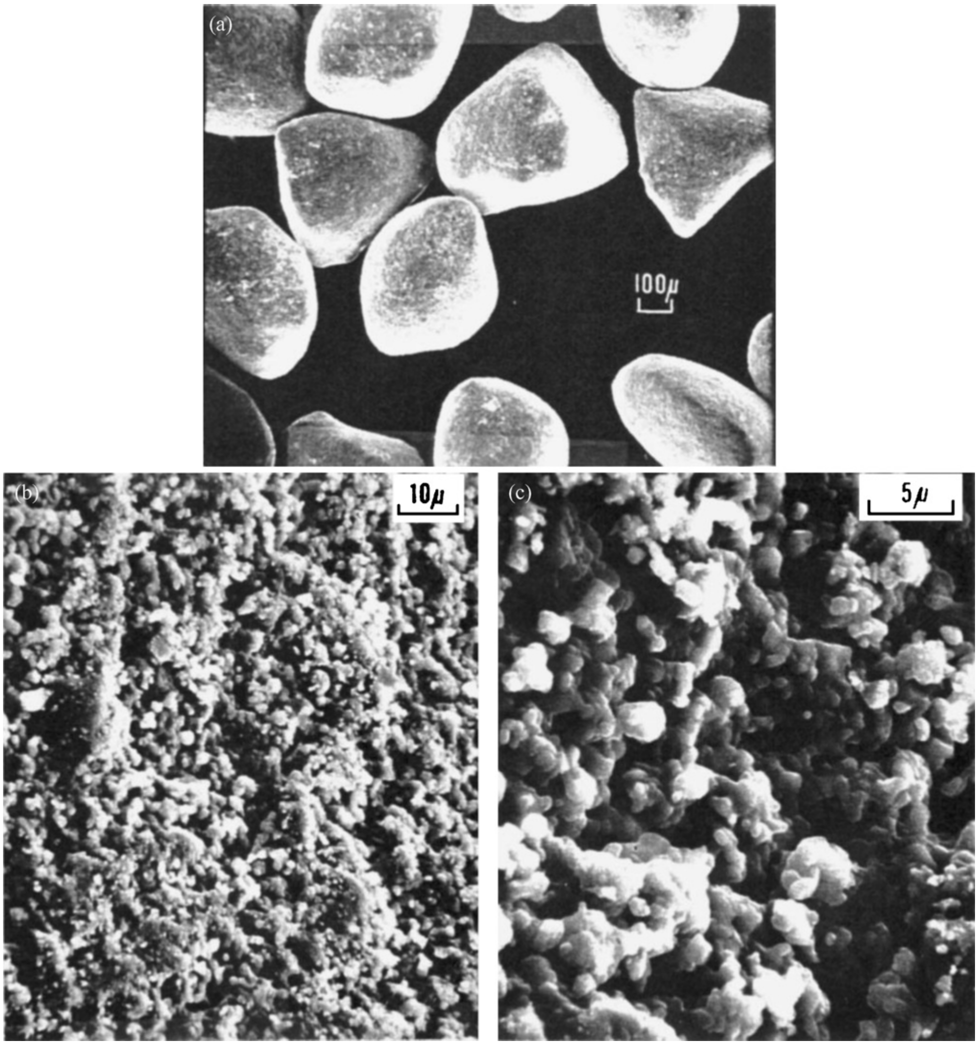


Figure 11
(a) SEM micrograph of a 12 grain size sintered bauxite or sintered alumina showing rounded surface with no sharp cutting edges. (b and c) SEM micrographs of sintered bauxite (or sintered alumina) at higher magnifications showing details of the knobby structure generated by intergranular separation in the green state or during subsequent sintering. Individual particles are seen to be diffusion bonded together into clusters and separated by microvoids (Komanduri and Shaw 1974a).

material. Magnesium nitrate hydrate, which converts to an MgO spinel during firing can be added to the sol-gel mix to inhibit grain growth. The fired material is cooled and separated to various grit sizes by sieving. Figure 12(a) is an SEM micrograph of a tough sol-gel abrasive and Fig. 12(b) is an SEM micrograph at higher magnification showing the microcrystalline nature of this abrasive responsible for high toughness at the same time sharp cutting action generating good finish.

Sol-gel abrasive described above involves the production of desired grit size by comminution of larger particles and grading the product to various grit sizes by sieving. This approach can be somewhat wasteful since all grit sizes may not be useful and as such material may have to be discarded, recycled, or used for low value applications. To overcome this problem, an interesting development in the synthesis of sol-gel abrasive took place which involves the formation of sol-gel alumina filaments, with an aspect ratio in the range of 2:1 to 8:1 which requires no or minimal comminution. The green abrasive filaments are formed from the gel by a variety of techniques including extrusion and spinning. Extrusion is the

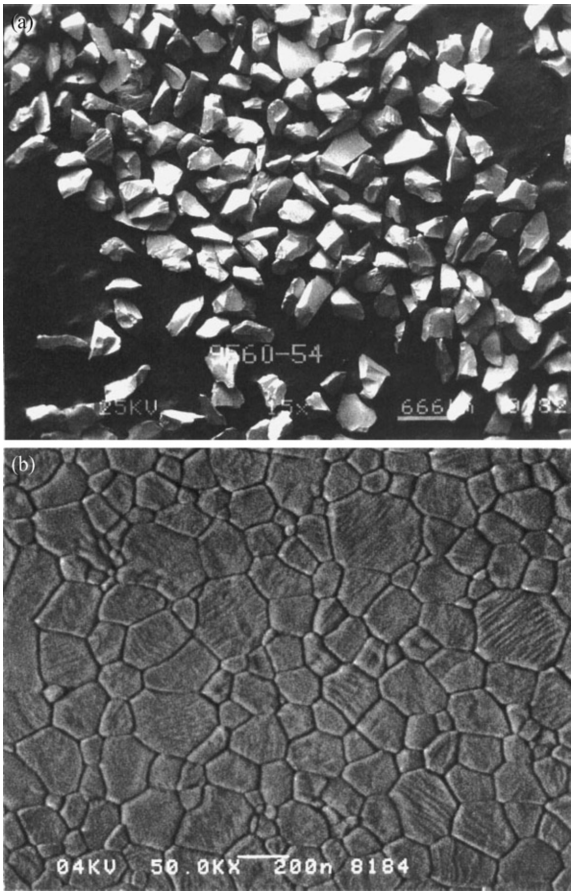


Figure 12
(a) SEM micrograph of a 46 grit tough, sol-gel aluminum oxide abrasive and (b) SEM micrograph at higher magnification showing the microcrystalline nature of the sol-gel aluminum oxide abrasive.

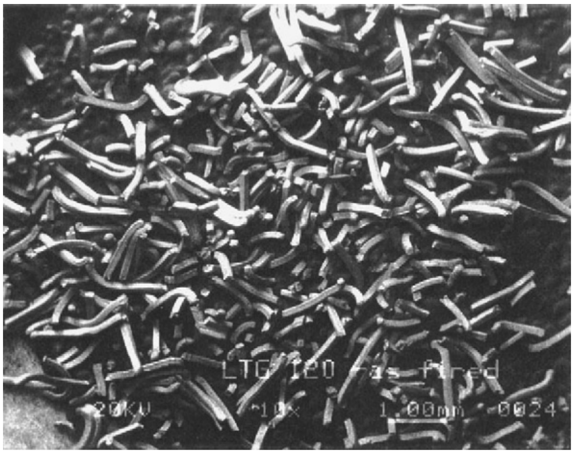


Figure 13
SEM micrograph of a filamentary sol-gel aluminum oxide abrasive.

preferred mode for green filaments between 0.25–1.5 mm in diameter, which after drying and firing, are roughly in the range of 100–24 grit size. Spinning is preferred for fired filaments $<100\,\mu m$. Fired filaments as fine as $0.1\,\mu m$ can be produced by this method. For the filaments formed by extrusion, the seeded gel is extruded through a smooth walled die with multiple holes to produce continuous gel filaments. They were then dried for 24–72 hours at a temperature of 75–80 °C and a relative humidity of $>85\%$. The filaments are converted to fibrous bodies with an average length of 2–8 mm. They were then converted to α-alumina by heating at a rate of $<2\,^{\circ}C\,min^{-1}$ to 800 °C, at a rate of $\sim 5\,^{\circ}C\,min^{-1}$ to 1370 °C, held at this temperature for ~5 minutes and then allowed to cool. For a gel filament of ~1.19 mm diameter, the filaments after firing are ~0.58 mm diameter corresponding to ~30 grit size. For filaments formed by spinning, ~1–5% of a nonglass forming spinning aid, such as polyethylene oxide is added to the gel to impart the desirable viscoplastic properties for the filament formation. Figure 13 is an SEM micrograph of a sol-gel alumina filamentary material produced by a special extrusion process.

Alumina–zirconia is a simple eutectic system with a very limited solid solution at higher temperature (Fig. 14). When zirconia is added to alumina, the structure becomes more equiaxial. Addition of ~40% zirconia to alumina results in the eutectic composition. Three compositions commercially considered for this alloy are 10% zirconia, 25% zirconia, and 40% zirconia. It is preferable to use the zirconia in the tetragonal form. By rapidly quenching on thin plates, the high temperature stable tetragonal form is retained instead of reversion to the weaker monoclinic form. As the zirconia content increases, the amount of eutectic structure increases resulting in an increase in the toughness at the expense of hardness. Thus, a 10% zirconia–alumina alloy is harder than 25% or 40% alloy with the toughness inversely proportional. This abrasive is manufactured by the fusion of a mixture of calcined bauxite, zircon sand, petroleum coke, iron filings, and rare earth modifiers in an electric arc furnace at ~1900 °C. The modifiers serve to stabilize the zirconia component of the abrasive grain.

The molten material is then poured in thin layers on a water cooled steel container or carbon clad receptacles containing calcium chloride at a temperature of 800 °C. This ensures rapid cooling of the molten material to a temperature in the range of 1350–1550 °C. This is the temperature range within which the crystal growth stops in alumina–zirconia alloy. Further cooling of this rapidly cooled material

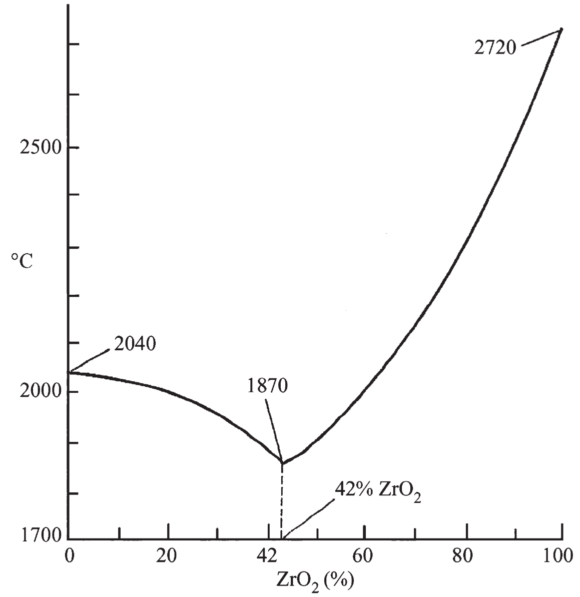


Figure 14
Al_2O_3–ZrO_2 phase equilibrium diagram (after Coes 1971).

to a temperature in the range of 500–800 °C proceeds slowly at a rate of 20 °C min^{-1}. This slow cooling helps in producing tensile stress-free and crack free crystals. The rapidly solidified and fine grain structure of alumina–zirconia alloy provides high impact strength and toughness which are desirable characteristics for an abrasive to be useful for heavy grinding applications, such as snagging, conditioning, and cut-off operations. The eutectic structure is composed of zirconia dendrites in an α-alumina matrix. Figure 15(a) shows an SEM micrograph of a 46 grit tough alumina-40% zirconia abrasive and Figure 15(b) is an SEM micrograph higher magnification showing the eutectic microstructure.

4. Synthetic Abrasives—Silicon Carbide

The crystal structure of α-silicon carbide (SiC) is hexagonal ($a = 3.073$ Å and $c = 15.07$ Å and $c/a = 4.899$). The structure of SiC is based on the stacking of hexagonal Si–C layers such that each Si atom is tetrahedrally surrounded by four C atoms and each C atom is also tetrahedrally surrounded by four Si atoms. The decomposition temperature of SiC is in the range of 2000 °C to 2400 °C and the Knoop hardness of SiC is ~2550 kg mm^{-2}.

In the synthesis of SiC, a mixture of silica sand, coke, and salt (to convert impurities to volatile chlorides) is charged into a furnace. The furnace comprises of a trough-like base with removable sides made of steel frame to which are fastened refractory bricks. At the ends are concrete or brick "heads" through which carbon or graphite electrodes pass (Fig. 16). The electrodes are connected to an electrical supply. The charge in the furnace surrounds the graphite core. Typical processing conditions include an applied voltage ~200–500 V, 60 Hz current with a power input of ~1800 kW, and operating at a total power of 54 MWH. The time under power is ~30 hours. Such a furnace is typically ~15 m long and produces an ingot of SiC ~8100 kg.

As the electrodes are powered, an electric current flows through the core causing it to become very hot (~2600 °C). The mixture adjacent to the core begins to convert into SiC according to the sequential reactions: $SiO_2 + 2C \rightarrow Si + 2\ CO$; $Si + C \rightarrow SiC$; and $SiO_2 + 3C \rightarrow SiC + 2CO$. The temperature range for good SiC formation is ~1800–2200 °C. The carbon monoxide is permeated through the reactive mix and burns to CO_2 at the surface of the mix in the furnace as a soft, rippling blanket of blue flame. SiC by itself is an electrical conductor.

As the mixture around the core is converted to SiC, it begins to consume some of the electric current, necessitating a downward adjustment in the applied voltage to limit the power input. After the run, the furnace is cooled, the sides removed, and the excess partially reacted mix is stored separately for reuse. The ingot is broken apart and sorted out manually. Since, considerable thermal gradient exits outwardly of the core, the ingot is not uniform in structure but consists of concentric zone of varying quality.

Various crushing and shaping stages are employed to crust the SiC ingots, and the product may be chemically treated with alkali to remove free silicon, and with acid to remove iron. Washing it with water containing surface active agents help in removing any free graphite present in the material. The grain is customarily roasted in a rotary kiln to condition the surface and remove any traces of organic matter. Grading to standard grit sizes is done in stages using screens.

Two types of SiC abrasive are marketed commercially, namely, the green SiC and the black SiC. The green variety is more pure than the black. The difference in color stems from the raw materials used and the conditions used in the furnace during firing. Green SiC is produced using a new mix each time while unreacted material from the previous mix is recirculated in the synthesis of black SiC. The abrasive performance of the two types is similar and the small improvement in hardness and performance is offset by higher cost of green SiC. Bulk of the SiC used as abrasive is black SiC. Figure 17 is an SEM micrograph of a 12 grit green SiC abrasive showing porosity due to volatile CO gases formed during manufacture of the abrasive. The surface features are similar to that of a fracture glass indicating very brittle nature of this abrasive. The high friability and low strength of SiC is attributed to these voids and

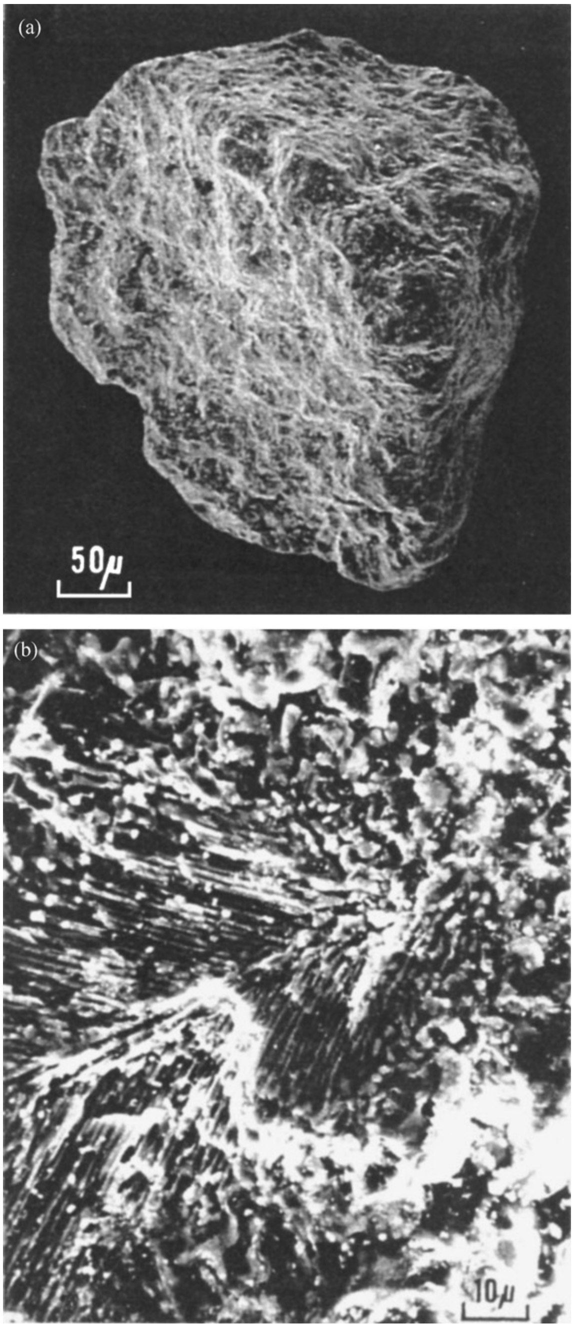


Figure 15
(a) SEM micrograph of a 50 grain size 40% zirconia alloyed alumina abrasive showing a tough blocky-type structure with no porosity and (b) SEM micrograph of a typical 40% zirconia alloyed alumina abrasive at higher magnification showing eutectic microstructure (Komanduri and Shaw 1974a).

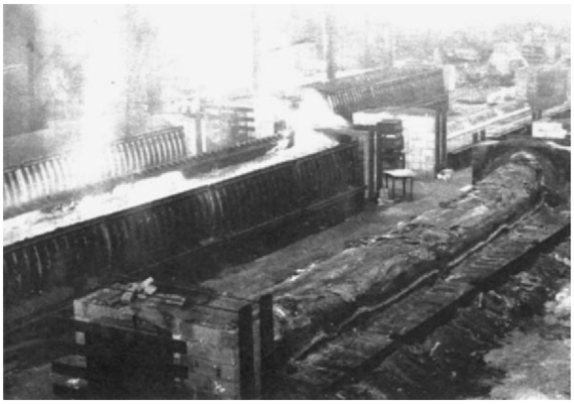


Figure 16
Photograph of an Acheson-type furnace used for the synthesis of SiC from SiO_2 and carbon.

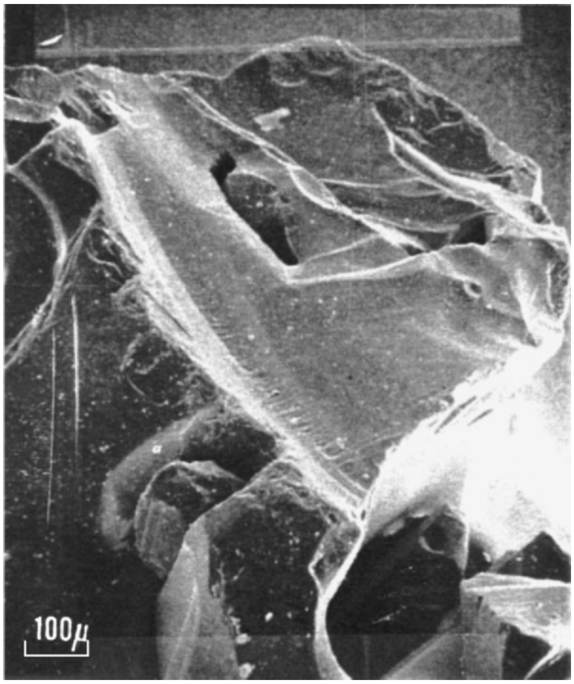


Figure 17
SEM micrograph of a 12 grain size green silicon carbide abrasive showing several holes formed as a result of the evolution of carbon monoxide during the manufacture with surface features similar to that of a fractured glass, indicating brittleness of the material (Komanduri and Shaw 1974a).

the relatively weak bonding between carbide particles joined to form a grain.

Boron carbide (B_4C) is a very hard abrasive (Knoop hardness $\sim 3400\,\mathrm{kg\,mm^{-2}}$), next only to diamond (till the synthesis of cBN) in hardness. It is used extensively in loose or powered form for polishing and as an abrasive in ultrasonic machining. It is, however, not used in grinding as an abrasive due to rapid wheel wear associated with the oxidation of B_4C to B_2O_3. Boron carbide is commercially manufactured by initially heating boric acid in an electric furnace to remove water of crystallization and to form anhydrous boric oxide, B_2O_3. To this is added high purity petroleum coke and the charge is heated in an electrical resistance furnace to about 2400 °C. Sawdust is added to provide pores for the escape of large quantities of reducing gas generated. Boron is condensed from the vapor phase and combines with carbon to form crystalline boron carbide.

5. *Synthetic Abrasives—Superabrasives*

In the materials world, the successful synthesis of diamond and CBN in the mid-1950s can be hailed as one of the most important technological and scientific achievements of the twentieth century. Diamond (Knoop hardness $\sim 8000\,\mathrm{kg\,mm^{-2}}$) and cubic boron nitride (cBN) (Knoop hardness $\sim 4700\,\mathrm{kg\,mm^{-2}}$), the two hardest materials known, come under this category and are termed the superabrasives. The synthesis of diamond has not only enabled the economic manufacture of diamond for industrial use but also facilitated tailoring of grain size, grain morphology, strength, and friability by controlling the pressure, temperature, time, and other processing conditions during the synthesis. It may be noted that grains of different strength and friability are required for different applications. It has also addressed the strategic aspect of this material. The direct production of diamond of a given size eliminates the need for crushing which weakens the product leaving behind cracks of uncontrolled size.

Long before the successful synthesis of diamond, the thermodynamic pressure–temperature equilibrium line between diamond (the cubic form of carbon) and graphite (the hexagonal form of carbon) was postulated by Rossini and Jessup of the National Bureau of Standards (now NIST) in 1938. Much of the initial high pressure work related to diamond synthesis was performed by Professor Bridgman of Harvard University before World War II. He was able to maintain a temperature of ~3000 K at a pressure of ~29 kbar but only for a short interval of time, conditions not sufficient for the direct conversion of graphite into diamond. It remained for a team of innovative researchers at General Electric (GE) to discover that certain catalyst–solvents made it possible to increase the rate of conversion to a practical level at relatively low temperatures (~1500 °C) and pressures (~50 kbar).

Hall (then of GE) built a HP-HT press, known as the "belt," and accomplished the successful synthesis of diamond along with his co-researchers. Figure 18(a) is a temperature–pressure equilibrium diagram showing diamond growth region in the presence of nickel as a catalyst–solvent and Fig. 18(b) is a schematic of the "belt" HP-HT press.

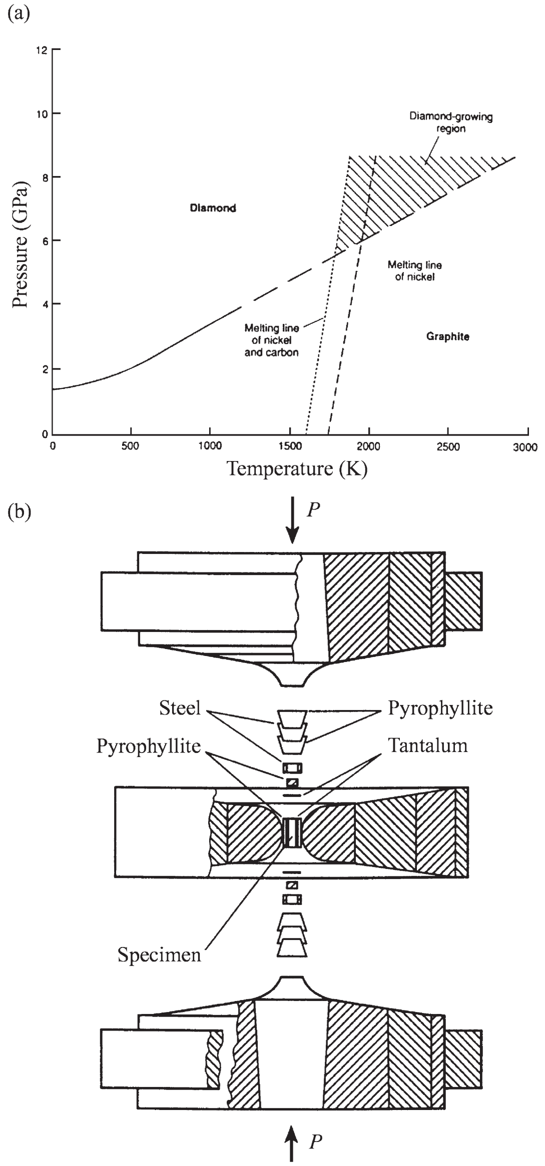


Figure 18
(a) Pressure–temperature equilibrium phase diagram showing the stability regions of diamond and graphite and the region of diamond growth in the presence of nickel as a catalyst–solvent (Clark and De Vries) and (b) HP-HT press, know as the "belt," used in the synthesis of diamond and cBN (after Hall 1960).

The elements that are effective as a catalyst–solvent include the transition elements Cr, Mn, and Ta, plus all the elements of group III of the periodic table. Carbides and compounds of these elements that decompose at or below the diamond synthesis conditions can also be used. These elements and their compounds are believed to play a duel role, namely, that of a catalyst, and as a good solvent for graphite. It appears that graphite first dissolves into the catalyst and it then converts into diamond at the appropriate conditions of pressure and temperature within the diamond stable region. Diamond being relatively insoluble in the molten catalyst, precipitates thus allowing more nondiamond form carbon to go into solution. Since, the catalyst–solvent must be in the molten state, this establishes a lower temperature limit for a given catalyst.

The rate of diamond nucleation is sensitive to pressure, the higher the pressure above the equilibrium line at a given temperature, the greater the rate of nucleation and growth. Diamond formed at pressures substantially above the equilibrium line develop from many nuclei and have a skeletal structure. Such a diamond is very friable and is designated as RVG, since they are used in resin or vitrified bonded grinding wheels for grinding cemented tungsten carbide. Conversely, by subjecting the reaction mixture to pressures and temperatures close to the equilibrium line for a longer time, fewer nucleation sites develop and larger and more perfect single crystals of diamond are formed. Temperature has a major influence on the crystal habit of the diamond produced. At the lowest temperatures, the cubic form is predominant. Mixed cubes, cubo-octahedra, and dodecahedra are preferred at intermediate temperatures while octahedra from at the highest temperatures. The color of the diamond produced also vary with the temperature from black at low temperatures, through dark green, light green, and yellow at intermediate temperatures, to white at the highest temperatures. The color pretty much depends on the nitrogen content and graphite inclusions. Since the lowest possible temperatures at which conversion into diamond takes place are controlled by the melting temperature of the catalyst metal–carbon eutectic system, it appears advantageous to use performed alloys that can reduce the minimum temperatures. This has the additional advantage of simultaneously lowering the pressure requirements. Lower temperatures and pressures can increase the life of the dies used and hence the economics of the process.

The most common metal alloys used with graphite, are Fe–Co–Ni (Mn) although other catalysts can be used. The effectiveness of nickel could be due to the fact that the unit cell of nickel is very close to that of diamond. In one process, alternate layers of metal alloy and graphite are loaded in the high-pressure press. First, the pressure is raised to the diamond stable region. Following this, the temperature is raised to the diamond stable region. They are held at these conditions for the required amount of time to complete the synthesis. Following it, the cell is cooled maintaining the high pressure and then the pressure is released. The rate of nucleation of diamond under HP-HT conditions is estimated to be $\sim 10000\,\mu\text{m}\,\text{min}^{-1}$.

In the lower temperature region but at pressures exceeding that of diamond–graphite equilibrium by a few kilobars, the rate of nucleation and growth of diamond is very high. The growing crystals interfere and produce a skeletal structure with numerous defects. The resulting grain is relatively weak and friable and is designated as RVG by one of the producers of diamond abrasive. Such a material is ideal for making resin-bonded grinding wheels to grind cemented tungsten carbide without causing thermal damage to the surface. Figures 19(a) and (b) are SEM micrographs of RVG diamonds (50–60 grain size) of single grain polycrystalline diamond at higher magnification showing the skeletal structure. Diamonds of this type are self-sharpening and provide sharp cutting edges without forming wear flats and rubbing. Although microchipping wear may be considerable, it frequently is justified in terms of improved surface integrity.

The life of the polycrystalline diamond abrasives of the skeletal structure can be improved significantly by encapsulating them in a metal sheath, chiefly Ni or Cu. This is important because fine grains of diamond or cBN can get dislodged before they can be worn and resharpened because of weak interfacial bonding between the abrasive and the resin bond. The metal coating add texture to the surface of diamond or cBN and retain the abrasive in the resin bond by improved mechanical anchoring. Typically, electroless plating is used for this purpose. Originally a thickness corresponding to $\sim 55\%$ of the weight of diamond (which corresponds to an increase in mean diameter of $\sim 15\%$) was used. In addition to providing more surface area of the grains for improved bonding, the metal sheath acts as a shield and strengthens the polycrystalline diamond. The metal also improve the heat transfer from the diamond abrasive to the surroundings. Nickel encapsulation on the diamond abrasive is used for resin bonded diamond grinding wheel for wet grinding while copper is used for dry grinding. Figures 20(a) and (b) are SEM micrographs of Ni and Cu plated polycrystalline diamond showing details of the surface morphology of the metal coating. Unlike the copper coating that conforms to the topography of polycrystalline diamond, the nickel coating completely encompasses the grain and the topographical features of the base crystal are not discernible. Sometimes, to increase the bonding of the nickel coated crystals to the resin bond of the grinding wheel, the material is subjected to further heat treatment that provides a spiky surface (Fig. 20(c)). The thickness of metal encapsulation was varied from 30–70% depending on the applications. For example,

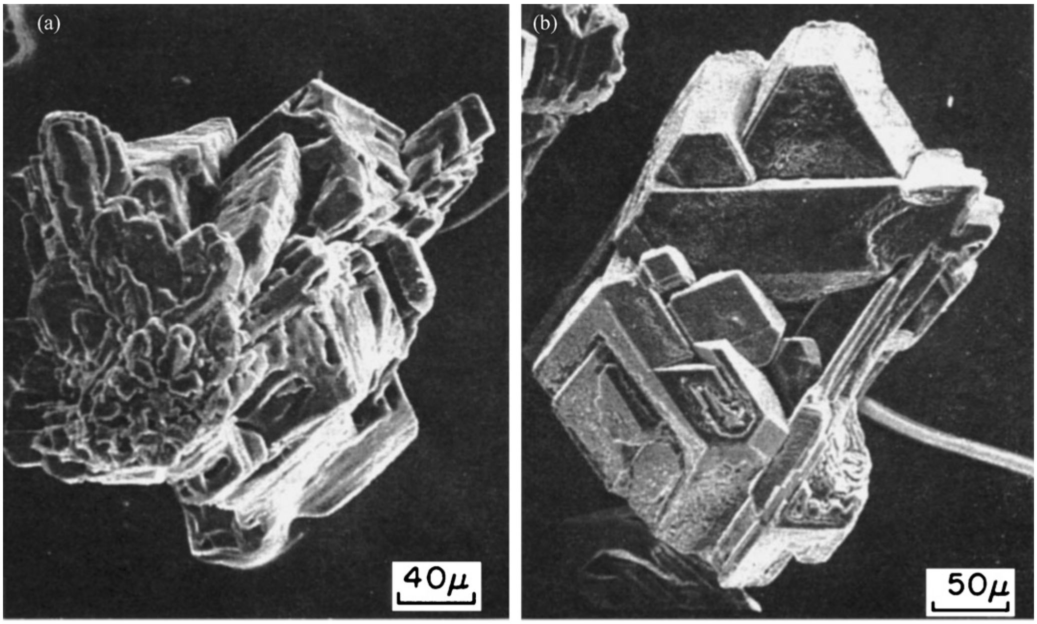


Figure 19
SEM micrographs of polycrystalline diamond crystals showing (a) and (b) skeletal structure (Komanduri and Shaw 1974a).

for obtaining longer wheel life in resin bonded grinding wheels, thicker coatings (~70%) are used to minimize premature dislodgement of the abrasives from the wheel while for applications requiring lower power requirements, thinner coatings (~30%) are used. Also, for metal bonded grinding wheels, the abrasives are coated initially with a thin layer of either Ti or Cr to enhance the bonding between the superabrasive and the metal bond.

Polycrystalline diamond of the RVG-type can also be grown using the low pressure–high temperature (LP-HT) CVD diamond synthesis, although the rates of growth are much lower (~150 μm) than the HP-HT technology. Figure 21 is an SEM micrograph of polycrystalline diamond grown using combustion synthesis by LP-HT synthesis. Polycrystalline diamond or cBN cutting tools made by the HP-HT synthesis can also be reused as polycrystalline diamond or cBN by breaking up the diamond layer. However, its use as a commercial product would depend on the economics of making such an abrasive by this route. While the strength of the polycrystalline diamond (or cBN) by this method is significantly higher, the cost can also be high, if prepared exclusively for this purpose. However, this may be one way to reclaim used polycrystalline diamond (or cBN) cutting tools for further use. Of course, if this method is used, then the polycrystalline diamond (or cBN) layer has to be produced without the underlying cemented carbide substrate, unlike the current practice of producing diamond (or cBN) cutting tool with a cemented carbide substrate.

Single crystal diamond of the MBG and MBS are stronger than the polycrystalline diamond (RVG-type). They are used generally in the metal bonded grinding wheels or saw blades for cutting, such hard materials as stone, concrete, and ceramics. The principal crystal habit of growth of these crystals is cubo-octahedral, although some faces may not be well formed. There are several varieties of MBG, namely, MBG II, MBG-P, and MBG-T. Due to a wide range of imperfections in the structure these crystals have medium friability and strength and fall between the weak RVG and the fully developed blocky MBS-type. Figure 22 is an SEM micrograph of a fully developed, cubo-octahedral, single crystal diamond formed under favorable growth conditions. MBG-T crystals, though fairly well developed, differ from other crystals in that their surfaces are etched and are uniformly rougher on a microscale. This will provide better bonding between the diamond crystals and the metal in the metal bonded diamond grinding wheels. Also, a thin coating of Cr or Ti is deposited on MBS crystals to enhance adhesion between the crystals and the metal binder used in the grinding wheels or saw blades.

Cubic boron nitride (cBN), next only to diamond in hardness (Knoop hardness 46.1 GPa, ~4700 kgf mm^{-2}), was developed in the late 1960s by Robert Wentorf Jr, of the GE. It is a remarkable material in

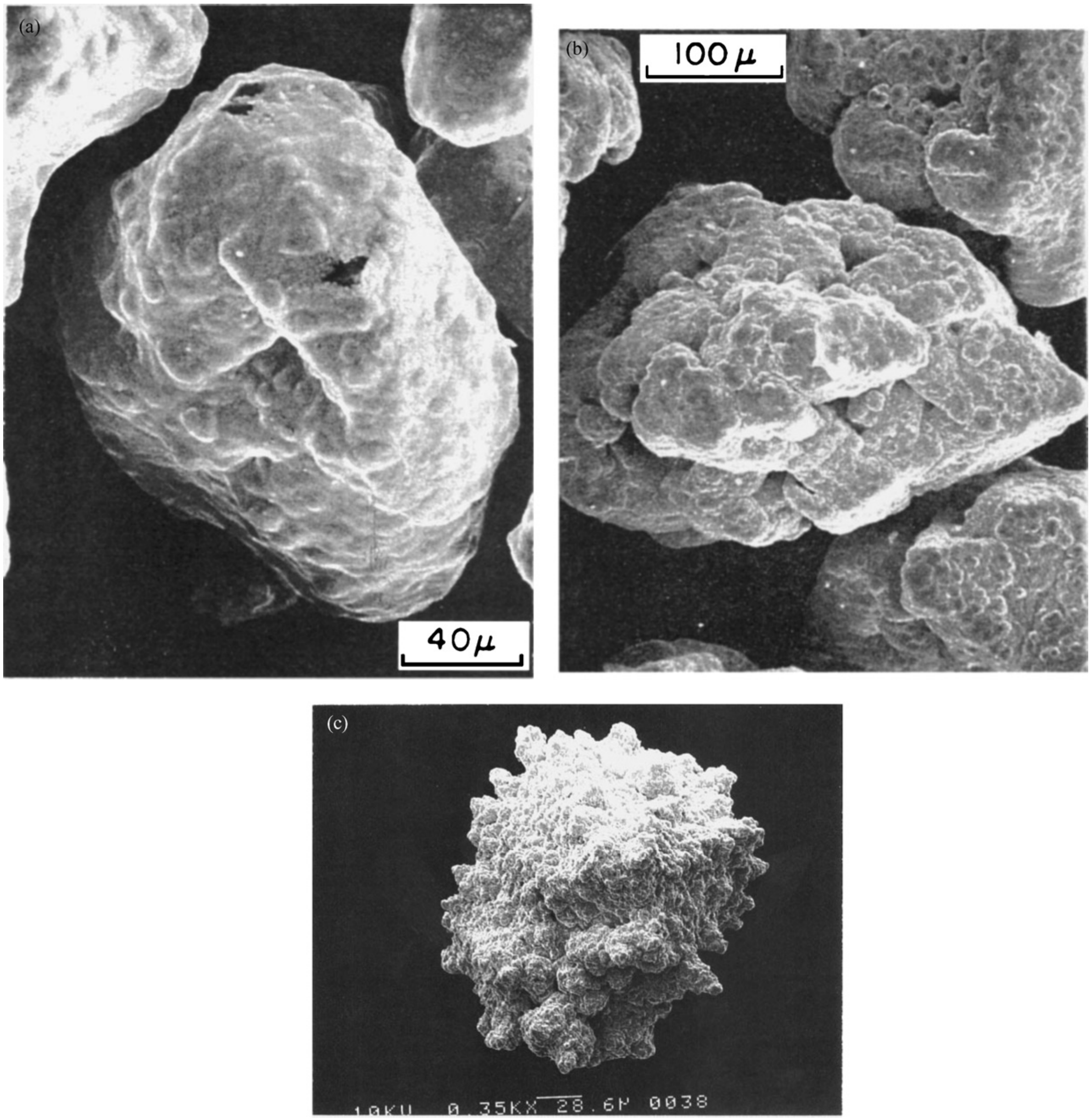


Figure 20
SEM micrograph of a nickel coated polycrystalline diamond crystal showing complete encapsulation of the crystal by the coating used for wet grinding of hardened steels. (b) SEM micrographs of a copper coated polycrystalline diamond crystal used for dry grinding of hardened steels. Note that the metal conforms to the grain topography and the character of the base polycrystalline diamond is maintained. (c) SEM micrograph of a spicked nickel coating on a polycrystalline diamond crystal to increase the bonding of the crystal in the resin bond and to improve the grinding ratio significantly in wet grinding of hardened steels (Komanduri and Shaw 1974b).

that it does not exist in nature and is produced by the high temperature–high pressure (HP-HT) synthesis in a process similar to that used to produce diamond from graphite. Hexagonal boron nitride (hBN) is used as the starting material. Alkali and alkaline-earth metal and their compounds (instead of Ni in the case of diamond) are found to be suitable catalyst–solvents for the production of cBN by the HP-HT process.

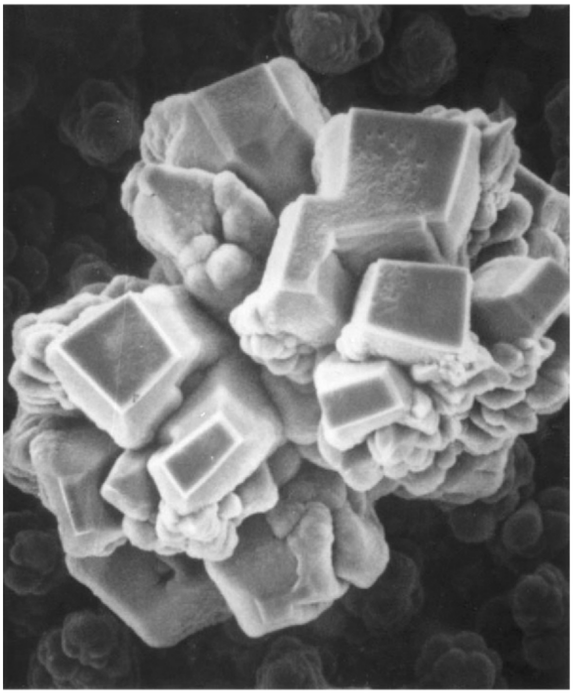


Figure 21
SEM micrograph of a polycrystalline diamond crystal grown by the LP-HT CVD diamond technique.

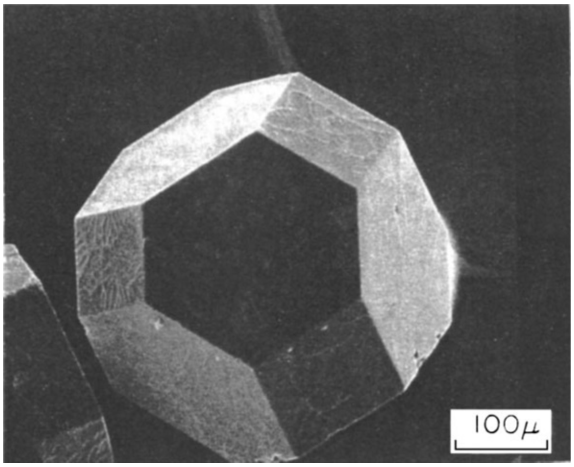


Figure 22
SEM micrographs of well formed cubo-octahedral single crystal diamond crystals used in saw blades (see Komanduri and Shaw (1974b).

In 1981, a tough, microcrystalline cBN abrasive was developed based on the direct conversion of cBN from pyrolitic hexagonal BN (hBN) without the use of a catalyst–solvent. This process involves pretreatment of hBN in a vacuum furnace to remove the boric oxide from the surface and form a layer of free boron followed by HP-HT synthesis. Initially hBN powder is loaded in a vacuum furnace and the furnace is evacuated to pressures of $\sim 1\times10^{-5}$ to 1×10^{-6} torr. It is subsequently heated to a temperature of 1400–1900 °C for a period of 60–90 minutes. The pressure increases to $>1\times10^{-3}$ torr during heating due to the evolution of nitrogen gas as a result of surface decomposition of the hBN power. This is followed by cooling of the material under vacuum, removing the hBN powder from the vacuum furnace and loading it on to the HP-HT press. Synthesis of cBN is carried out for about 3–10 minutes at a pressure of ~55–80 kbar and a temperature of ~1600 °C in the absence of any catalyst–solvents as well as impurities (which would interfere with the conversion or sintering of cBN). Such a material (with or without a metal coating) being very tough (amongst various types of cBN) can be used for heavy material removal operations and at the same time can be used for obtaining good surface finish due to the microcrystalline nature of the abrasive.

Cubic boron nitride, though not as hard as diamond, is less reactive with ferrous materials, such as hardened alloy steels, hard chill-cast iron, and nickel-base and cobalt-base superalloys. Hence, it can be used to grind these materials. cBN crystallites are stronger than polycrystalline diamond of the RVG-type but weaker than fully formed, single crystal (MBS-type) diamond. The principal growth habits of cBN crystals are octahedra and tetrahedra. Figure 23(a) is an SEM micrograph of a cBN crystal showing a well developed crystal with sharp edges. The surfaces of these crystals are somewhat smooth and featureless. At high growth rates twinning is predominant. Figure 23(b) is an SEM micrograph of a twinned cBN crystal. Figure 23(c) is an SEM micrograph of a microcrystalline cBN crystal grown without a catalyst–solvent. Similar to the coatings on RVG diamond, nickel is also used for coating cBN crystals (Fig. 24).

6. Synthetic Abrasives—Chemo-mechanical Abrasives

In mechanical finishing, where an abrasive harder than the workmaterial is used, the surface damage due to abrasion can be minimized but not eliminated altogether even when a very fine, hard abrasive is used. To improve the final surface finish and to meet the application requirements for reliable performance in service, further finishing needs to be carried out preferably involving, what is known as chemo-mechanical polishing (CMP). With the advent of microelectronics circuits, CMP has assumed a much broader role in that it not only encompasses polishing but also extends it to produce plane surfaces or planarization. So, the ‘P’ in the ‘CMP’ is extended to include planarization.

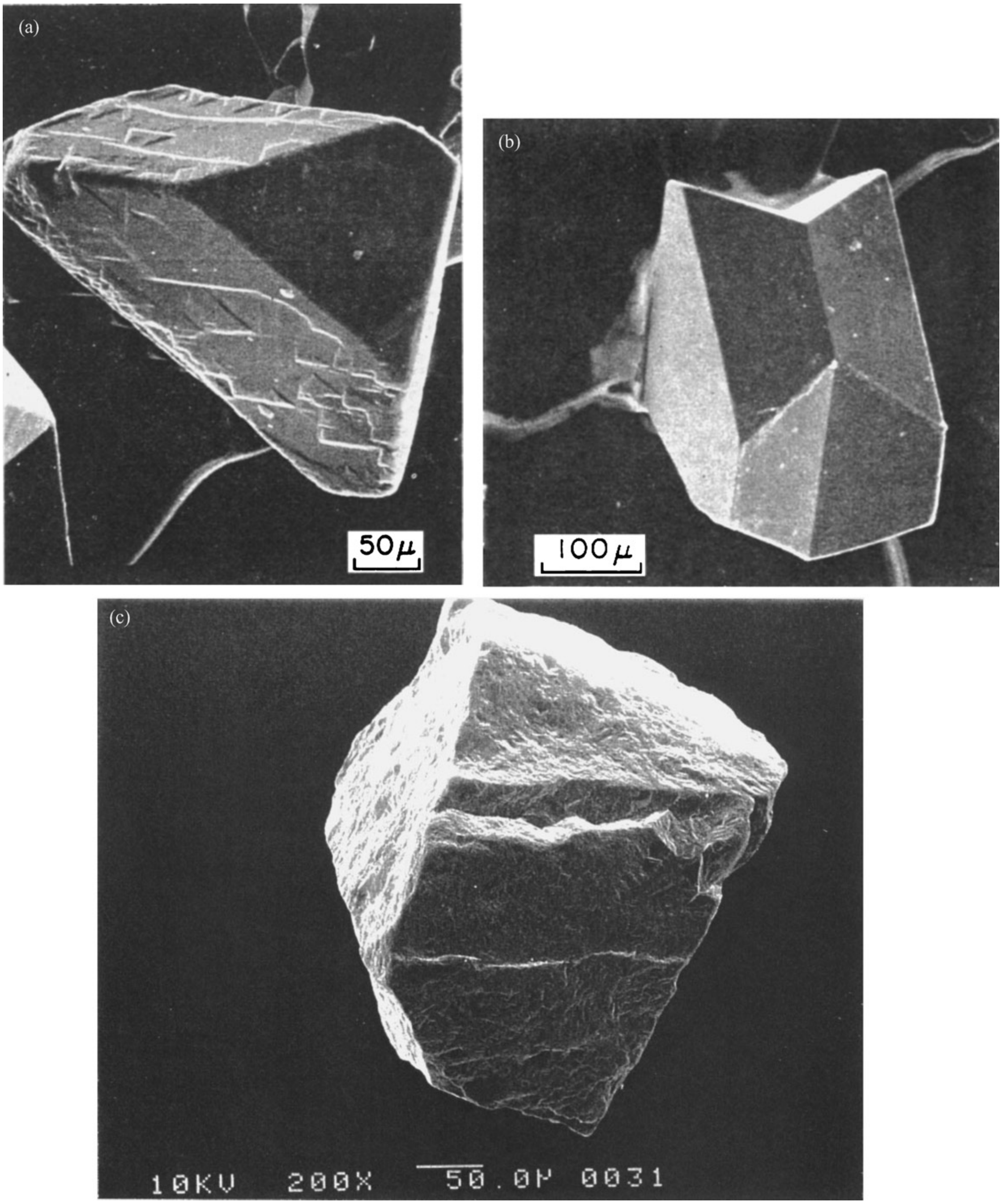


Figure 23
(a) SEM micrographs of a single crystal cubic boron nitride (cBN) indicating well developed planes and sharp edges. (b) SEM micrographs of a single crystal twinned cubic boron nitride (cBN) crystal. Note the relatively smooth and featureless surface. (c) SEM micrograph of a tough, microcrystalline cubic boron nitride (cBN), obtained using higher pressures and temperatures and in the absence of any catalyst–solvent normally used (Komanduri and Shaw 1974b).

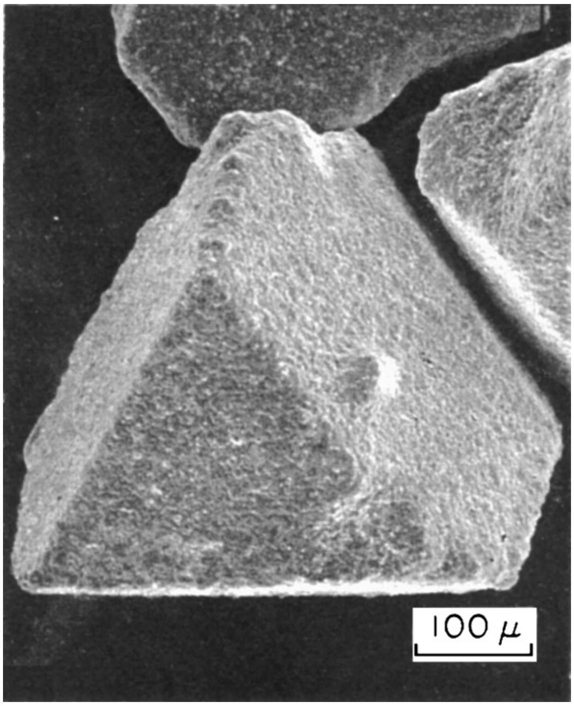


Figure 24
SEM micrographs of a nickel coated cubic boron nitride (cBN) crystal showing complete encapsulation of the crystal by the coating used for wet grinding of hardened steels (Komanduri and Shaw 1974b).

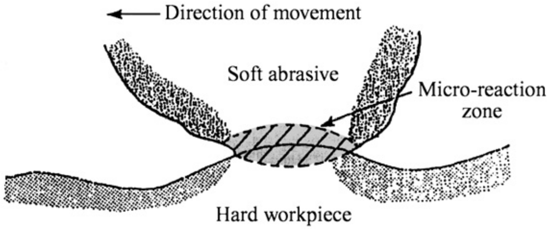


Figure 25
Schematic showing chemo-mechanical action between the abrasive, the workmaterial, and the environment (after Yasunaga *et al.* 1979).

CMP depends on the availability for a short duration of a certain threshold pressure and temperature at the contact zone of the polishing process provided a chemical reaction layer can be formed by the interaction of the workmaterial, the abrasive, and the environment (Fig. 25). The material removal by CMP is believed to be at the molecular level thereby yielding a smooth surface. The formation of the reaction layer and its removal controls the process. Similarly, the dissolution of the reaction products in the carrier fluid also plays an important role.

Many abrasives softer than the workmaterial can be used for finishing some of the difficult-to-finish workmaterials, such as advanced ceramics, glasses, and silicon. Here, material removal occurs due to chemical interaction between specific combinations of the abrasive–workmaterial–environment under the conditions of operation, such as the local pressure and temperature at the interface. Abrasives that come under this category include cerium oxide (CeO_2), magnesium oxide (MgO), silica (SiO_2), chromium oxide (Cr_2O_3), alumina (Al_2O_3), titania (TiO_2), zirconia (ZrO_2), iron oxide, or rouge (Fe_3O_4). For example, it has been found that damage-free surfaces can be produced using silica (SiO_2) for polishing sapphire; $BaCO_3$, CeO_2, and $CaCO_3$ for polishing silicon single crystals; iron oxide or rouge (Fe_3O_4) and magnesia (MgO) for polishing quartz; Cr_2O_3, CeO_2, and ZrO_2 for finishing Si_3N_4. In the microelectronics industry, nanoparticle-sized SiO_2 is used for oxide polishing, while alumina and silica are used for metal polishing. CeO_2, MgO, TiO_2, ZrO_2, iron oxide or rouge (Fe_3O_4), and hafnia (HfO_2) are also used for polishing.

Ceria or cerium dioxide (CeO_2) has a cubic crystal structure with a fluorite-type structure while the cerous oxide (Ce_2O_3) has a hexagonal structure. They are rare earth compounds. The melting point of ceria is ~2750 °C. Although cerium occurs in nature in its trivalent form, cerium dioxide (CeO_2) is more stable than cerous oxide (Ce_2O_3), particularly in the presence of free oxygen. There are two valence forms of cerium, namely, the trivalent (cerous) as in Ce_2O_3, and the tetravalent (ceric) as in CeO_2. Cerium oxide is produced using cerium salt solutions. Cerium carbonates or hydroxides are precipitated and then calcined in rotary kilns or muffle furnaces to cerium oxide at temperatures of ~1000 °C. The impurities (CaO, SrO, nonrare-earth elements) must not exceed certain level in the raw material; otherwise, cerium oxide will agglomerate during calcination and lose polishing capability. The calcined concentrates are milled and cleaned by sieving.

Cerium dioxide, or ceria (CeO_2) is produced by electrolytic oxidation of cerium (III) to cerium (IV) ions, or chlorination with pure chlorine gas. Ceria salts are easily hydrolyzable and precipitate in the form of insoluble basic salts, leaving cerous salts in solution. The calcination of basic salts results in the formation of ceria as a fine white power. Ceria is extensively used for the polishing of optical glass and has been found to be a superior polishing agent than iron oxide rouge. It is also found to be an excellent polishing agent for CMP of silicon nitride for reasons similar to that of glass polishing. Its use can also be extended for the finishing of silicon as well. CeO_2 has been reported to be particularly effective in water and in alcohols only when hydroxyl groups are present. Thus CeO_2 slurry in water is required for polishing.

The oxide contains polyvalent cerium atoms, Ce(IV) and Ce(III), which can provide chemical action with the workmaterial and the environment.

Magnesium oxide or magnesia (MgO) is a very soft material with Mohs hardness of ~5.5 which is in between apatite(s) and orthoclase(s). The melting point of magnesia is ~2800 °C. The principal commercial source of magnesia is magnesite ($MgCO_3$) and dolomite ($MgCO_3 \cdot CaCO_3$). Magnesia is produced by calcining dolomite and treating the resulting mixture of magnesia and calcia with a strong solution of ammonium chloride, or with a solution of magnesium and calcium chlorides. The insoluble magnesia is then separated from the calcium chloride solution. Magnesia is used as an abrasive in CMP.

Chromium oxide (Cr_2O_3) has a hexagonal rhombohedral structure of the corundum-type. Cr_2O_3 is green in color when in a finely dispersed form. It melts at 2435 °C and has hardness slightly higher than alumina. Hence, it can be used singly as an abrasive like alumina or alloyed with alumina. The production of Cr_2O_3 involves the reduction of sodium dichromate, generally with sulfur. The mixture is fed into a brick-lined furnace, brought to a dark-red heat when the reaction proceeds exothermically. After cooling, the sodium sulfate produced is leached out with water. The remaining solid is separated, rinsed, dried, and ground.

Cr_2O_3 is a well-known catalyst. It is this role of Cr_2O_3 that is widely believed to take in the polishing. However, recent studies have shown that the role of Cr_2O_3 is more than that of a catalyst, at least in the polishing of Si_3N_4. It is found that Cr_2O_3 does participate in the CMP of Si_3N_4 to form chromium nitrate (CrN) and chromium silicate (Cr_2SiO_4), in addition to the formation of a silica layer on Si_3N_4. Cr_2O_3 is widely used as a chemo-mechanical abrasive. However, some of the reaction products are not environmentally friendly and hence its use will in future be somewhat limited till this problem has been solved.

Titania (TiO_2) is found in nature in the form of rutile. It is manufactured mainly from ilmenite ($FeTiO_3$), a widely available mineral. It is leached with sulfuric acid and the hydroxide subsequently precipitated from the solution is then calcined to TiO_2.

Zirconia (ZrO_2) is produced from zircon sand ($ZrO_2 \cdot SiO_2$). At room temperature its structure is monoclinic, which changes to tetragonal at higher temperatures, and to cubic at still higher temperatures. It is used in the natural form for abrasive jet cutting and as a coated abrasive. Pure zirconia is used as a polishing medium or alloyed with alumina to form alumina–zirconia for stock removal grinding (SRG) applications where toughness of the eutectic mixture plays an important role.

Silica or silicon dioxide (SiO_2) is available in crystalline, nano to microcrystalline, and noncrystalline forms. Silica constitutes ~60% of the earth's crust either alone or combined with other oxides or silicates. Silica represents a limiting case in which an infinite 3D network is formed by the sharing of all the oxygen atoms of a given tetrahedron with neighboring groupings. For microelectronics applications, silica sols and colloidal silica are used. Colloidal silica is a stable aqueous dispersion or sol of discrete amorphous silica particles having a grain size of 1–100 nm. The pH value should be ~7 to maintain sufficient negative charge on the silica particle surface to prevent aggregation. To produce sols that are stable at relatively high silica concentrations, particles must be grown to a certain size in a weakly alkaline aqueous dispersion.

Iron oxide or rouge (Fe_3O_4) has been used for a long time for the polishing of glass, and precious metals like silver and gold. In some cases it has been replaced by zirconia or ceria because they are cleaner materials to use and have less tendency to strain.

Bibliography

Bhagavatula S, Komanduri R 1996 On the chemo-mechanical polishing of silicon nitride with chromium oxide abrasive. *Philo. Mag. A* **74**, 1003

Bovenkerk H P, Bundy F P, Hall H T, Strong H M, Wentorf R H Jr 1955 Preparation of diamond. *Nature* **184**, 1094–8

Bradbury J C, Murry H H, Lefond S J 1986 Abrasives. In: Bever M (ed). Encyclopedia of Materials Science and Engineering. Pergamon, Oxford, UK, Vol. 1, pp. 13–24.

Bundy F P, Hall H T, Strong H M, Wentorf R H Jr. 1955 Man-made diamond. *Nature* 176.

Clark T J, DeVries R C 1990 Superabrsives and ultrahard tool materials. In: Lampman S R, Zorc T B (eds.) *Metal Handbook—Properties and Selection*, Vol. 2, 10th edn. pp. 1008–18

Coes L Jr 1971 *Abrasives*. Springer-Verlag, New York

Corrigan F R 1981 *US Pat*. 4 289 503.

Cottringer T E, van de Merwe R H, Bauer R 1996 *US Pat*. 4 623 364.

Garg A K, Khaund A P, Orne L E, Young M R 1998 *US Pat*. 5 725 162.

Hall H T 1960 *US Pat*. 2 947 610.

Hall H T 1980 Ultra-high pressure, high temperature apparatus: the 'belt'. *Rev. Sci. Instrum.* **31**, 125–31

Jiang M, Komanduri R 1998 On the chemo-mechanical polishing (CMP) of silicon nitride (Si_3N_4) workmaterial with various abrasives. *Wear* **220**, 59–71

Khaund A K, Garg A K, Bauer R 1993 *US Pat*.. 5 593 468.

Kilbourn B T 1992 Cerium—a Guide to its Role in Chemical Technology. Molycorp, White Plains, New York

Komanduri R, Shaw M C 1974a Scanning electron microscope study of surface characteristics of abrasive materials. *Trans. ASME J. Eng. Mater. Technol.* **96**, 145–56

Komanduri R, Shaw M C 1974bb Surface morphology of synthetic diamonds and cubic boron nitride. *Int. J. Mach. Tool Des. Res.* **14**, 63–84

Komanduri R, Nandyal S 1993 Low-pressure synthesis of polycrystalline diamond abrasive. *Int. J. Mech. Tools Manufact.* **33**, 285–96

Malkin S 1989 *Grinding Technology—Theory and Applications of Machining with Abrasives*. Ellis Horwood, Chichester, UK

Reu C V 1991 Abrasive. In: Howe- Grant M (ed.) *Kirk-Othmer Encyclopedia of Chemical Technology*, 4h edn. Wiley, New York, Vol. 1, pp. 12–37

Rue C V, van de Merwe R H, Bauer R, Pellow S W, Cottinger T E, Klok R J 1993 *US Pat.* 5 194 072.

Ryshkewitch E 1960 *Oxide Ceramics—Physical Chemistry and Technology*. Academic Press, New York

Shaw M C 1996 *Principles of Abrasive Processes*. Oxford University Press, Oxford

Ueltz H F G 1972 Abrasive grains—past, present, and future. In: Shaw M C (ed.) *Proc. Int. Grinding Conf.* Carnegie-Mellon University, Pittsburgh, PA, USA, pp. 1–52

Yasunaga N, Tarumi N, Obara A, Imanaka O 1979 Mechanism and application of the mechanochemical polishing method using softer powder. In: Hockey B H, Rice R W (eds.) *Science of Ceramic machining and Surface Finishing II* NBS Special Publication No 562 p. 171

R. Komanduri[a] and S. Iyengar[b]
[a] *Mechanical & Aerospace Engineering, Oklahoma State University, Stillwater, Oklahoma, USA*
[b] *Saint-Gobain/Higgins Grinding Technology Center, Worcester, Massachusetts, USA*

Creep and Creep-fatigue of Metal-Matrix Composites

A metal-matrix composite (MMC) comprises a metallic matrix reinforced by a ceramic phase, which can be added as continuous fibers, discontinuous fibers, or particulates. Some MMCs possess good mechanical properties at high temperatures and these MMCs are considered as potential candidate materials for advanced high-temperature components. High-temperature components operate in the creep range where they have to withstand static (creep) and cyclic loads (fatigue). The term creep-fatigue interaction is used to describe the complex service conditions that are encountered in high-temperature systems, which are characterized by stationary as well as periodically, or irregularly, changing loads.

1. Creep of Metal-matrix Composites

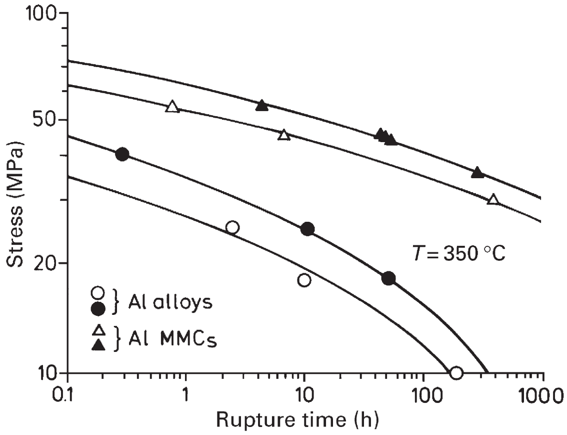


Figure 1
Stress–rupture curves of aluminum alloys with and without fibers (after Dlouhy *et al.* 1993).

Creep is the time-dependent plastic deformation of materials, which is important at temperatures above half of the melting point of a metallic material in kelvin (Riedel 1987). Creep strongly depends on microstructure, stress, and temperature. During tensile creep, strains accumulate until rupture occurs. Figure 1 illustrates how reinforcing an aluminum alloy with 15vol.% of Al_2O_3 fibers improves the creep rupture strength (Dlouhy *et al.* 1993). MMCs generally exhibit lower rupture ductilities than the matrix materials without reinforcements. Creep deformation of metals and alloys is generally characterized by a homogeneous type of deformation, which occurs in the whole volume of the specimen. Creep damage is generally an internal damage like creep cavitation, e.g., in copper or in ferritic steels (Riedel 1987). Creep data for complex engineering alloys like MMCs are commonly represented by equations, which originally were developed to describe the behavior of materials that had much simpler microstructures. For example, it is often assumed that a regime of steady state creep can be identified where the constant steady state creep rate $\dot{\varepsilon}_S$ depends on stress σ and temperature T as

$$\dot{\varepsilon}_S = C\sigma^n \exp\left(-\frac{Q_{\text{eff}}}{kT}\right) \qquad (1)$$

where C is a constant, n is the stress exponent, and Q_{eff} is the apparent activation energy of creep, and k is the Boltzmann constant. MMCs differ in creep behavior from pure metals and simple alloys in two important respects: (i) there is no steady-state creep regime, but an inflection from primary to tertiary creep resulting in a minimum creep rate; and (ii) the use of Eqn. (1) to represent minimum creep rate data leads to values of n and Q_{eff} that are much larger than those associated with creep of simple metals and alloys. For short fiber-reinforced aluminum alloys the shape of individual creep curves as well as the stress and the temperature dependence of the minimum creep rate could be rationalized on the basis of an interaction of three elementary microstructural processes, which are schematically illustrated in Fig. 2. Fibers are loaded by dislocations, which form a work hardened zone around them. This results in a stress transfer from the matrix to the fibers. Recovery processes in the course of which dislocations from the work hardened zone move to the fiber ends (recovery path) where they shrink and annihilate counteract

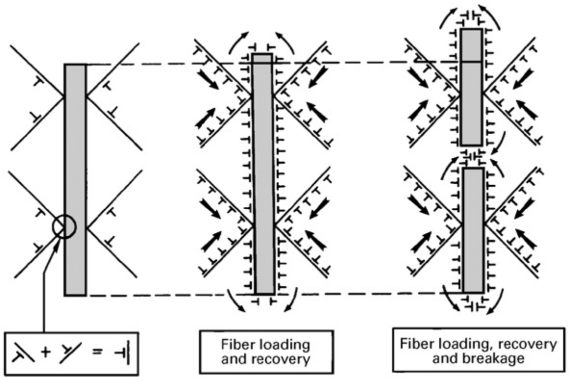

Fiber loading and recovery
Fiber loading, recovery and breakage

Figure 2
Elementary microstructural processes which govern creep of short fiber-reinforced aluminum alloys (after Dlouhy *et al.* 1995).

this strain hardening process. Once fibers reach a critical stress they break and this reduces the length of the recovery path. The example given here (Dlouhy *et al.* 1995) describes the creep behavior of one specific type of short fiber-reinforced MMC. However, the results are of general validity in showing that it is possible to rationalize high values of Q_{eff} and n (Eqn. (1)) on the basis of the interaction of elementary deformation processes, which can be detected in the MMC microstructure.

2. *High-temperature Fatigue of MMCs and Creep-fatigue Interaction*

Fatigue is characterized by the failure of a material subjected to cyclic loading (in stress or strain control) which occurs after a given number of load cycles. One of the load cycles alone would not result in failure. In fatigue, damage accumulates with increasing cycle number, as creep damage does with increasing time. Plastic deformation under fatigue conditions is often associated with localized plastic deformation. Fatigue life is generally associated with the initiation and growth of surface cracks and strongly depends on the environment. In high-temperature fatigue experiments, the response of a material to cyclic loading is investigated at creep temperatures. This has important consequences for the mechanisms of fatigue deformation and damage accumulation. High-frequency fatigue experiments are often found to yield results, which are quite similar to what is observed in low-temperature fatigue tests. In low-frequency experiments there is more time for recovery and this results in a decrease of the saturation stress observed in the hysteresis of a total strain controlled low cycle fatigue (LCF) test. There is also more time for an interaction between the environment and surface cracks. The first effect is often considered as representing the influence of creep, the second effect represents the influence of high-temperature corrosion on fatigue life. With respect to a microstructural understanding of high-temperature fatigue and creep-fatigue three questions need to be answered: (i) what effect has creep-fatigue loading on the stress strain hysteresis; (ii) how do microscopic cracks develop in the surface of the MMC (nucleation of engineering cracks); and (iii) how does a macroscopic fatigue crack propagate in the microstructure of an MMC?

In total strain controlled isothermal testing, creep-fatigue interaction can be addressed in various ways, such as: (i) determining the influence of the imposed total deformation rate; (ii) studying the effect of different wave shapes; (iii) investigating the effects of hold times in tension or compression on fatigue life; (iv) low cycle fatigue testing using pre-crept specimens; or (v) performing creep and fatigue tests sequentially. It is well known that high-temperature fatigue loading can be anisothermal such that temperatures and stresses change simultaneously, "in- and out-of-phase." Such types of loadings are referred to as thermal mechanical fatigue (TMF). There are different experimental and theoretical approaches to high-temperature fatigue and to creep-fatigue interaction of MMCs ranging from TMF tests to modeling creep-fatigue interaction on the basis of constrained cavity growth models. In the case of tungsten fiber-reinforced copper it was shown that TMF loading resulted in lower fatigue lives than isothermal LCF. In the case of *in situ* composite solders it was not possible to fully account for creep-fatigue interaction merely on the basis of a constrained cavity growth model. From a microstructural point of view it has been qualitatively discussed how the superposition of static and cyclic high-temperature loading in a MMC affects dislocation dynamics near the fiber-matrix interface (Eggeler and Earthman 1997). There is little information on which factors affect high-temperature fatigue crack initiation in MMCs.

3. *The Role of the Reinforcement-matrix Interface, Mechanical Anisotropy, and Multiaxial Loading*

Given the wide range of different types of MMCs and the multitude of possible creep-fatigue loading conditions the present text cannot provide one simple set of clear microstructural conclusions on creep and high-temperature fatigue behavior of MMCs. However, in both, creep and fatigue of MMCs the role of the interfaces between reinforcement and matrix is crucial. Strengthening of MMCs relies on load transfer from the matrix to the fiber across the interface. In the case of fiber-reinforced composites, fracture toughness is influenced by fiber pullout events and crack propagation is affected by the interaction of the propagating crack with the fiber-matrix interface. Ductility is affected by relaxation of stresses near the fiber-matrix interface. Mechanical anisotropies also are important. Thus a short fiber-reinforced MMC with aligned fibers can be strong at high temperatures in the fiber direction where fibers can constrain the material during creep and toughen the material during fatigue crack growth (where fiber pullout can dissipate energy). On the other hand, when loaded perpendicular to the fibers, debonding of the reinforcement-matrix interface under creep conditions will result in poor creep properties. Fast fatigue crack propagation along a reinforcement-matrix interface—which would also be expected under low-temperature fatigue conditions (Suresh 1991)—will result in poor fatigue performance. High-temperature components generally have to withstand multiaxial loading and this adds a further degree of complexity (Leckie 1996).

4. Conclusions

It is more difficult to design high-temperature components using MMCs. Their properties are dependent on the mix of matrix and reinforcement materials, on the spatial distribution of the constituent phases, on the properties of fibers and reinforcements, and on the build up of the fiber-matrix interface. The high-temperature properties of MMCs moreover depend on the interaction and coupling of elementary microstructural deformation and damage processes. A better understanding of all aspects of high-temperature loading of MMCs is required to enable safe design of high-temperature MMC components.

Bibliography

Chawla K K 1998 *Composite Materials, Science, and Engineering*, 2nd edn. Springer, New York

Clyne T W, Withers P J 1993 *An Introduction to Metal-matrix Composites*. Cambridge University Press, Cambridge

Dlouhy A, Eggeler G, Merk N 1995 A micromechanical model for creep in short fiber-reinforced aluminum alloys. *Acta Metall. Mater*. **43**, 535–50

Dlouhy A, Merk N, Eggeler G 1993 A microstructural study of creep in short fiber reinforced aluminum alloys. *Acta Metall. Mater*. **41**, 3245–56

Eggeler G F, Earthman J C 1997 Dislocation fiber interactions in short fiber-reinforced aluminum alloys during creep and during thermal cycling. *Scr. Mater*. **38**, 341–8

Leckie F A 1996 High-temperature mechanism-based design. In: Cahn R W, Evans A G, McLean M (eds.) *High-temperature Structural Materials*. Chapman & Hall, London, pp. 193–205

Riedel H 1987 *Fracture at High Temperatures*. Springer, Berlin

Skelton R P (ed.) 1987 *High Temperature Fatigue: Properties and Predictions*. Elsevier Applied Science, London

Suresh S 1991 *Fatigue of Materials*. Cambridge University Press, Cambridge

Williams J C 1996 Material requirements for high-temperature structures in the 21st century. In: Cahn R W, Evans A G, McLean M (eds.) *High-temperature Structural Materials*. Chapman & Hall, London, pp. 17–31

G. F. Eggeler
Ruhr-Universität, Bochum Germany

Creep of Bituminous Aggregates

Bitumen is a complex hydrocarbon produced for industrial use by oil refining. For its various applications it is almost always mixed with mineral aggregate, the particular blend depending on the end use. The largest field of application is for roads and other paved areas but it is also used for waterproofing of roofs, bridge decks, and underground construction.

The oil industry is able to produce a wide range of bitumens, including many with properties modified through the use of a variety of polymers or other additives.

The principal characteristics of bitumen, which are attractive for its various applications, are its adhesive and waterproofing properties (see also *Bituminous Materials*[#]).

1. Basic Physical Properties of Bitumen

Bitumen is a viscous fluid at medium to high temperatures and a brittle solid at low temperatures. Its consistency is quantified as viscosity and as stiffness, which is defined as applied stress divided by corresponding strain, for these temperature regions, respectively. The relationship between consistency and temperature is a very important one, since the use of bitumen requires it to be in a fluid state for mixing with mineral aggregate and a somewhat stiffer state for compaction by roller to form a road layer and to act as a solid for its in-service role at ambient temperatures.

Bitumen is a viscoelastic material, the properties of which are influenced not only by temperature but also by rate of loading. Figure 1 shows the three phases of behavior including elastic, viscoelastic, and viscous as the temperature rises. It is the viscous property of bitumen that causes it to exhibit creep characteristics. These are apparent both when it is subjected to static loading and under the repeated loading which arises from moving traffic on highways or runways. Figure 2 shows the effect of deformation in the wheel tracks of a major highway causing a dangerous level of rutting.

Many grades of bitumen may be regarded as "rheologically simple" meaning that a given rise in temperature has the same effect on stiffness as a

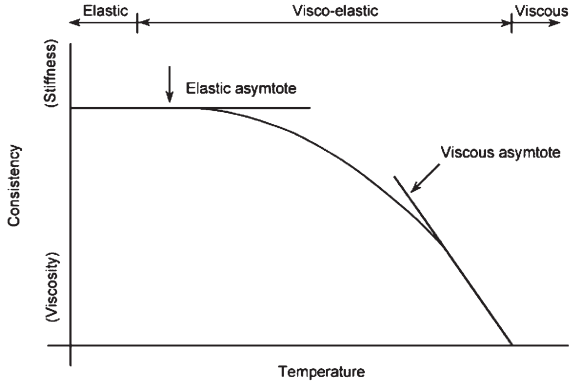


Figure 1
General relationship between consistency and temperature for a bitumen.

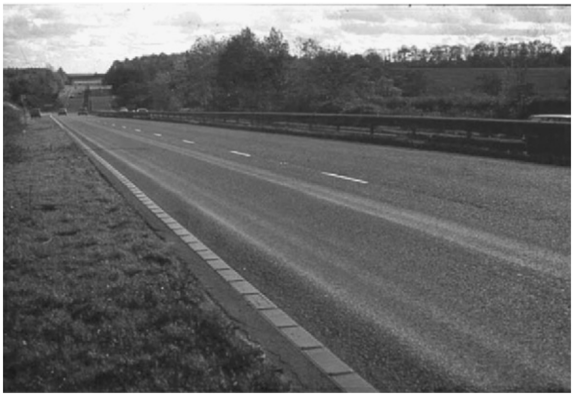


Figure 2
Serious rutting on a major highway.

corresponding decrease in rate of loading. It follows that the stiffness or viscosity of bitumen may be determined for a very wide range of temperature or loading rate by conducting experiments over a limited range and "shifting" the data to extrapolate over a large range to produce a "master curve" of response. The result of such a procedure is shown in Fig. 3 for both a normal bitumen and a polymer-modified material.

2. *Creep Characteristics of Bitumen–Aggregate Mixtures*

Bituminous material used for road construction has three phases, typical volumetric properties being 12% bitumen, 83% mineral aggregate, and 5% air voids. The creep characteristics of this mixture are influenced by the properties of the two solid components.

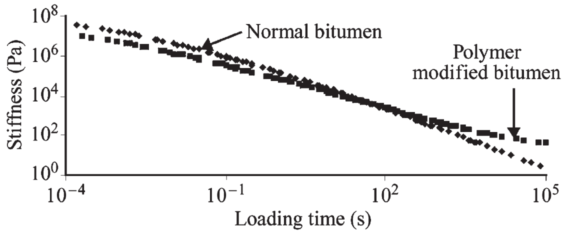


Figure 3
Master curves relating stiffness to loading time for a normal and an elastomer-modified bitumen.

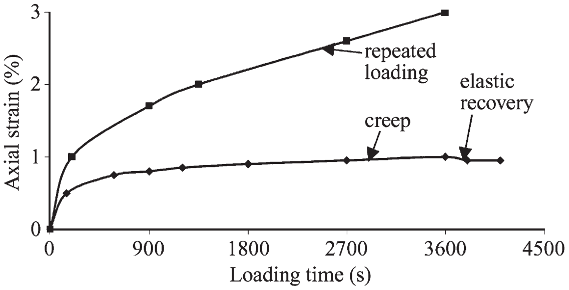


Figure 4
Relationships between strain and loading time for tests under static and repeated loading on a typical bituminous material.

However, the presence of air is important, since an aggregate matrix saturated with bitumen would exhibit high creep deformation, which would be unacceptable in practice and cause serious rutting of the type shown in Fig. 2.

When a bituminous mixture is subjected to applied stress, there are, in general, three components of strain which result: elastic, plastic, and viscous. The plastic component is independent of time and temperature, being generated by the interaction between aggregate particles. Both the other components are time and temperature dependent. Elastic strains are influenced by the bitumen and the aggregate, while viscous strain derives only from the bitumen.

Under a static creep load, Fig. 4 shows how the total strain develops with time. When the load is removed, a small elastic (recoverable) strain is apparent. Under repeated loading of the same magnitude, the accumulated strains are larger. This is partly because of the plastic strains contributed by the relative movement between aggregate particles.

The addition of a polymer changes the consistency/temperature relationship for a bitumen as shown in Fig. 3 for an elastomer-modified material. When mixed with aggregate, this produces a material that will be less susceptible to temperature changes and more resistant to creep strains than one made with a conventional bitumen.

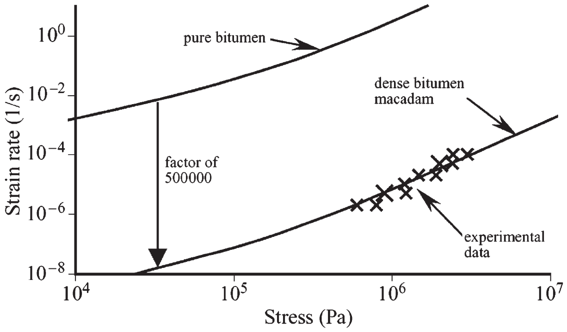


Figure 5
Relationships between strain rate and stress at 20 °C.

The effect of increasing the volume of aggregate in a bituminous mixture is shown in Fig. 5 for a typical case. These data show how the steady strain rate (slope of the linear parts of relationships in Fig. 4) varies with applied uniaxial stress and illustrates the substantial reduction in strain rate when aggregate is blended with bitumen. The relationship between these two parameters is only linear up to a particular stress level; in this case about 50 kPa. In the linear range the response is known as Newtonian with strain rate proportional to stress, while in the non-linear range the material is described as non-Newtonian. Since viscosity is the quotient of stress and strain rate, it follows that a Newtonian bitumen has a unique value of viscosity at a particular temperature regardless of the stress or strain level. In reality, under working conditions, most bitumens are often in the non-Newtonian range.

S. F. Brown
University of Nottingham, UK

Creep of Concrete

Under sustained load, concrete continues to deform. This phenomenon is called creep. Creep, along with instantaneous (or elastic) strain, is defined as the strain difference between loaded and load-free specimens. When exposed to drying, the load-free specimens exhibit drying shrinkage (typically reaching 0.0002 to 0.0004, and in poor concrete even 0.0008). After many years, the creep strain typically attains values 2 to 6 times larger than the initial elastic strain. After unloading, a partial creep recovery takes place. The creep at drying greatly exceeds the sum of shrinkage and the creep of a sealed specimen. The difference, called the drying creep (or Pickett effect), is a manifestation of a coupling between deformation and water content changes. Imbibition of water causes swelling (normally much less than the drying shrinkage) or, if concrete has dried, a partial reversal of shrinkage.

Drying shrinkage (Fig. 1 left and middle) is caused by compressive stresses in the microstructure balancing changes in the capillary tension and surface tension on pore walls, as well as the disjoining pressure in hindered adsorbed water layers. The chemical processes of Portland cement hydration cause autogeneous shrinkage, which is observed in sealed specimens and amounts to only about 5% of drying shrinkage (but much more in high-strength concretes). Creep is caused by bond ruptures (with restorations at adjacent sites) in the hardened Portland cement paste, which is strongly hydrophilic, having a disordered colloidal microstructure, porosity about 0.4 to 0.55, and an enormous internal surface area—about $500\,\mathrm{m^2\,cm^{-3}}$. Its main component is the tricalcium silicate hydrate gel ($3\,\mathrm{CaO}\cdot 2\,\mathrm{SiO_3}\cdot 3\,\mathrm{H_2O}$, in short $\mathrm{C_3S_2H_3}$); it forms crystalline sheets and needles of colloidal dimensions, weakly bound by van der Waals forces.

1. *Constitutive Law*

At service stresses in structures, which are less than about one-half of the strength, and in the absence of cracking, the creep strain depends on stress linearly. It may be fully characterized by the compliance function $J(t,t')$ (Fig. 1 right), defined as the strain ε at time (or age) t caused by a unit uniaxial stress $\sigma = 1$ applied at age t'. As t' increases, the creep diminishes. This phenomenon, called aging, causes J to depend not only on the time lag $t-t'$. At variable stress $\sigma(t)$, each stress increment $\mathrm{d}\sigma(t')$ applied at time t' produces strain history $\mathrm{d}\varepsilon(t) = J(t,t')\mathrm{d}\sigma(t')$. By the superposition principle (introduced by Boltzmann and Volterra), one gets the linear creep law of aging viscoelasticity:

$$\varepsilon(t) = \int_{t_1}^{t} J(t,t')\mathrm{d}\sigma(t') + \varepsilon^0(t) \qquad (1)$$

where ε^0 denotes the sum of shrinkage strain $\varepsilon_{\mathrm{sh}}$ and thermal strain (and also smeared cracking strain, if any). The integral is the Stieltjes integral, which admits histories $\sigma(t)$ with jumps; for time intervals with no jump, one may set $\mathrm{d}\sigma(t') = [\mathrm{d}\sigma(t')/\mathrm{d}t']\mathrm{d}t'$. When history $\varepsilon(t)$ is prescribed, then Eqn. (1) represents a Volterra integral equation for $\sigma(t)$, which is not integrable analytically for realistic forms of $J(t,t')$, although numerical integration is easy. The solution $\sigma(t)$ for strain $\varepsilon = 1$ imposed at age $\hat{t}$ (and for $\varepsilon^0 = 0$) is called the relaxation function, $R(t,\hat{t})$ (Fig. 1 right) [$R(t,\hat{t}) \approx 0.992/J(t,\hat{t}) - 0.115\psi/J(t,t-1\mathrm{day})$, $\psi = J(\bar{t},\hat{t})/J(t,\bar{t}) - 1$, $\bar{t} = (\hat{t}+t)/2$]. Generalizing Eqn. (1) according to the principle of superposition, and taking isotropy into account, yields a triaxial stress–strain relation, in which the shear and bulk compliance

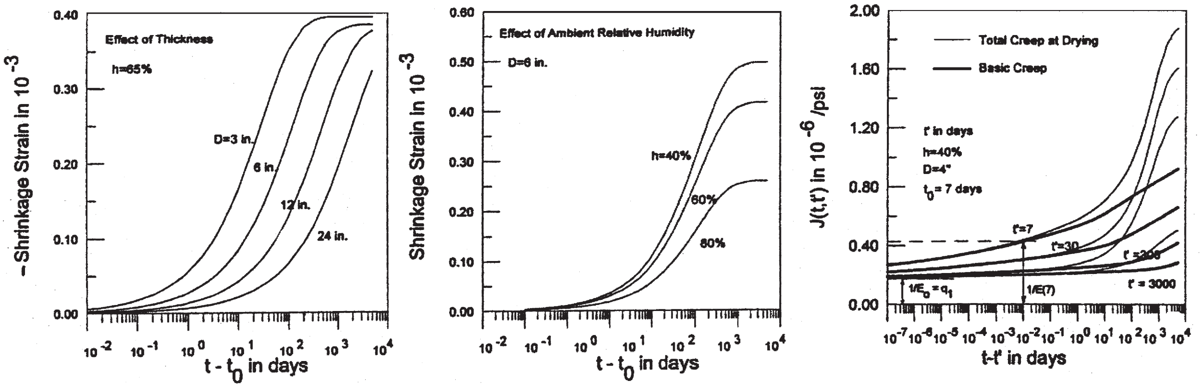


Figure 1
Typical time evolution of shrinkage, and of basic creep and drying creep compliances, for various slab thicknesses D, environmental humidities h_e, and ages at loading t'.

functions are $J_G(t,') = 2(1+2\nu)J(t,t')$ and $J_K(t,t') = 3(1-2\nu)J(t,t')$ ($\nu \approx 0.18$ = Poisson ratio). At high stress, the creep law appears to be nonlinear (Fig. 2) but Eqn. (1) remains valid if the strain due to time-dependent growth of cracks is included in $\varepsilon^0(t)$.

The value $J(t',t') = q_1$, corresponding to extrapolation of creep curves to zero load duration, may be considered as an age-independent constant. The conventional Young's elastic modulus, implied as $E(t') = 1/J(t'+\delta,t')$ where usually $\delta \in (0.0001\,\text{s}, 10\,\text{min.})$, increases with age t'. A realistic form of the compliance function (Bažant *et al.* 1997; bold curves in Fig. 1 right) may be conveniently expressed by its rate:

$$\dot{J}(t,t') = v^{-1}(t)\dot{C}_g(t-t') + 1/\eta_f \tag{2}$$

$$v^{-1}(t) = q_2(\lambda_0/t)^m + q_3 \tag{3}$$

$$\dot{C}_g(t,t') = \frac{n(t-t')^{n-1}}{\lambda_0^n + (t-t')^n} \tag{4}$$

For creep of sealed specimens, called the basic creep:

$$1/\eta_f = q_4/t(\text{at constant}\, w) \tag{5}$$

and then the flow part of $J(t,t')$ is simply $q_4\ln(t/t')$.

The superior dots denote partial derivatives with respect to t; θ = load duration; $\lambda_0 = 1$ day, $m = 0.5$, $n = 0.1$; q_2, q_3, q_4 = empirical constants; $C_g(t-t')$ = (age-independent) compliance function for delayed elasticity of the cement gel (hardened cement paste without its pores); $v(t)$ = volume of gel per unit volume of concrete, growing due to hydration; and η_f = effective viscosity for the flow of concrete. By integration, $C_g(t-t') = \ln\,\{1+[(t-t')/\lambda_0]^n\}$. As for $J(t,t')$, it can be obtained only by numerical integration (however, for computer structural analysis in time steps, the rate $\dot{J}(t,t')$ suffices and in fact allows a simpler algorithm).

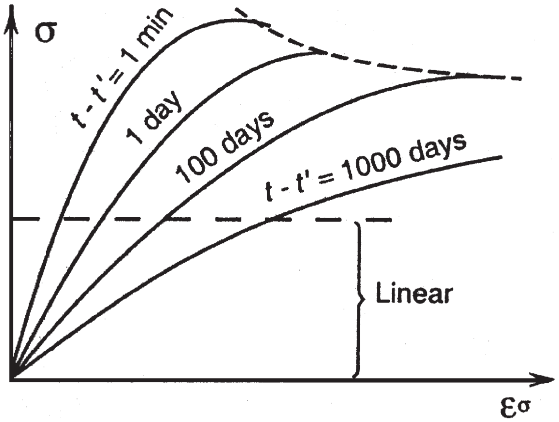


Figure 2
Creep isochrones.

2. *Variable Environment*

At variable w (mass of water per unit volume of concrete), a physically realistic constitutive relation may be based on the idea of microprestress S, considered to be a dimensionless measure of the stress peaks at the creep sites in the microstructure. The microprestress is produced by chemical volume changes and by changes in the disjoining pressures in the hindered adsorbed water layers (up to ten water molecules, or 2.7 nm, in thickness) confined between sheets of calcium silicate hydrates. The disjoining pressures must vary with the relative humidity h in the capillary pores, so as to maintain thermodynamic equilibrium (equality of chemical potentials). The rate of bond breakages may be assumed to be a quadratic function of the level of microprestress, which leads to the expression:

$$1/\eta_f = q_4 S \tag{6}$$

The microprestress is not appreciably affected by the applied load. It relaxes in time, and its evolution at each point of a concrete structure may be solved from the differential equation:

$$\dot{S} + c_0 S^2 = -c_1 \dot{h}/h \qquad (7)$$

where c_0, c_1 = dimensionless constants. The fact that changes of w or h produce new microstress peaks, and thus activate new creep sites, explains the drying creep effect (however, a part of this effect is caused by the fact that microcracking in a companion load-free specimen makes its overall shrinkage less than the shrinkage that occurs in a specimen creeping under compression, thus increasing the difference between the two).

The concept of microprestress is also needed to explain the stiffening due to aging. One physical cause of aging is that the hydration products gradually fill the pores of hardened cement paste, as reflected in function $v(t)$ in Eqn. (3). But hydration ceases after about one year, whereas the effect of age at loading t' is strong even after many years. The explanation is that the peak microstresses relax with age, which reduces the number of creep sites and the rate of bond breakages causing creep.

The evolution of distributions $h(\mathbf{x},t)$ ($\mathbf{x}$ = coordinate vector) may be considered uncoupled from the stress and deformation problem, and may be solved numerically from the diffusion equation $\dot{h} = \mathrm{div}[C(h)\ \mathrm{grad}\ h] + \dot{h}_s(t_e)$ where $h_s(t_e)$ = self-desiccation (caused by hydration) (which is mild in normal concrete but strong in high-strength concretes) and $C(h)$ = diffusivity, which decreases about 20 times as h drops from 100% to 60%. The field of free (unrestrained) shrinkage strain rates

$$\dot{\varepsilon}_{\mathrm{sh}} = k_{\mathrm{sh}} \dot{h} \qquad (8)$$

where k_{sh} = shrinkage coefficient. Since $\dot{\varepsilon}_{\mathrm{sh}}$ values at various points are incompatible, shrinkage stresses are induced, and the calculation of the overall shrinkage of structures as well as test specimens is a stress analysis problem, in which creep and cracking must be taken into account.

For finite element structural analysis in time steps, it is advantageous to convert the constitutive law to a rate-type form. This may be achieved by approximating $C_g(t,t')$ with a Kelvin chain model (or the associated relaxation function with a Maxwell chain model). The history integrals such as Eqn. (1) then disappear from the constitutive law, the history being characterized by the current values of internal state variables (the partial strains or stresses of the Maxwell or Kelvin chain).

Conversion to a rate-type form is also convenient to introduce the effect of variable temperature, which affects the Kelvin chain viscosities (according to Arrhenius law) and also the rate of hydration. The latter is captured by introducing into the laws for creep and shrinkage an equivalent hydration period, t_e, based again on the Arrhenius law for the hydration reaction. For a three-dimensional tensorial generalization of (2)–(8), see Bažant *et al.* (1997).

3. *Approximate Cross-section Response at Drying*

Although multidimensional finite element calculations of creep and moisture diffusion are nowadays feasible, simplified one-dimensional analysis of concrete beams or girders still reigns in practice. In that approach, one needs as input the average cross-sectional compliance function $\bar{J}(t,t',t_0)$ (Fig. 1 right, light lines) and average shrinkage function $\bar{\varepsilon}_{\mathrm{sh}}(t,t_0)$ of the cross-section (Fig. 1 left and middle) (t_0 = age at start of drying). The algebraic expressions for such average characteristics are considerably more complicated and inaccurate than the constitutive law for a material point, because of ignoring the differences due to cross section geometry, reinforcement, and loading (compressive or tensile axial force, bending moment, shear, torque, etc.). The following approximations (Bažant and Baweja 2000), partly based (under strong simplifications) on the foregoing constitutive relations, have been derived and their coefficients optimized by fitting a large test data bank; for environmental humidities h_e below 98%:

$$\bar{\varepsilon}_{\mathrm{sh}}(t,t_0) = -\varepsilon_{\mathrm{sh}\infty} k_{\mathrm{h}} S(t), \quad k_{\mathrm{h}} = 1 - h_{\mathrm{e}}^3 \qquad (9)$$

$$S(t) = \tanh\sqrt{\frac{t-t_0}{\tau_{\mathrm{sh}}}}, \quad \tau_{\mathrm{sh}} = k_{\mathrm{t}}(k_{\mathrm{s}}D)^2 \qquad (10)$$

where $D = 2v/s$ = effective thickness, v/s = volume-to-surface ratio, $k_t = 1$ for normal (type I) cement; k_s = shape factor (e.g., 1.0 for a slab, 1.15 for a cylinder); and $\varepsilon_{\mathrm{sh}\infty} = \varepsilon_{\mathrm{s}\infty} E(607)/(E(t_0 + \tau_{\mathrm{sh}})$, $\varepsilon_{\mathrm{s}\infty}$ = constant; $E(t) = E(28)\sqrt{4 + 0.85t}$ = age dependence of Young's modulus (all times are in days). Eqns. (2), (3) and (4) apply, except that $1/\eta_f$ must be replaced by:

$$\frac{1}{\bar{\eta}_f} = \frac{q_4}{t} + q_5 \frac{\partial}{\partial t}\sqrt{F(t) - F(t_0')} \qquad (11)$$

where q_4, q_5 are constants, $F(t) = \exp\{-8[1-(1-h_e)S(t)]\}$, and $t_0' = \max(t', t_0)$. The form of the expression for shrinkage halftime τ_{sh} is based on the diffusion theory. Function tanh in Eqn. (9) is the simplest function satisfying two asymptotic conditions ensuing from the diffusion theory: (i) for short times $\bar{\varepsilon}_{\mathrm{sh}} \propto \sqrt{t - t_0}$, and (ii) the final shrinkage must be approached exponentially. Generalizations taking into account the effect of temperature have also been formulated.

Empirical formulae have been developed for predicting the parameter values in the foregoing equations on the basis of concrete strength and some parameters of the concrete mix. However, they are very crude, leading to prediction errors with coefficients of variation 23% for creep and 34% for drying. These high uncertainties can be drastically reduced by updating certain coefficients of the formulae according to short-time creep and shrinkage tests of the given concrete. However, the evolution of weight loss of the drying test specimens must be measured simultaneously (or else the updating problem is ill-conditioned). A fully rational prediction of concrete creep and shrinkage properties from its composition is a formidable problem, not yet resolved satisfactorily.

4. Engineering Applications

Nonuniform creep and restrained shrinkage may cause deleterious cracking in bridges, buildings and other structures, which may trigger or promote other processes endangering durability. Creep nonuniformity is caused by differences in the histories of pore humidity and temperature, age, and concrete type. In slender structures, creep may cause collapse due to long-time instability. Transfer of compressive stresses from creeping and shrinking concrete to steel bars or other non-creeping parts (e.g., stone cladding or masonry liner) may induce their buckling or compression failure. The creep effects are particularly important for prestressed concrete structures (because of their slenderness and high flexibility), and are paramount in the safety analysis of nuclear reactor containments and vessels. The stress relaxation caused by creep is useful in mitigating the effects of shrinkage, cooling, differential creep, and differential settlements, but is undesirable in the case of a stress loss in prestressing tendons. At high-temperature exposure, as in fire or postulated nuclear reactor accidents, creep is very large and plays a major role.

Bibliography

Bažant Z P 1995 Creep and damage in concrete. In: Skalny J, Mindess S (eds.) *Materials Science of Concrete IV*, American Ceramic Socuety, Westerville, OH, pp. 355–389

Bažant Z.P., Baweja S 2000 Creep and shrinkage prediction model for analysis and design of concrete structures: Model B3. In: Special Publ. 194, Am. Concrete Institute, Farmington Hills, MI, pp. 1–83 (update of Bažant and Baweja, *Materials and Structures* (RILEM, Paris) **28**, 357–365, 415–430, 488–495; Errata, Vol. 29, p. 126)

Bažant Z P, Hauggaard A B, Baweja S, Ulm F-J 1997 Microprestress-solidification theory for concrete creep, Parts I and II. *J. Eng. Mech.* **123** (11), 1188–94 and 1195–1201

Bažant Z P, Kaplan M F 1996 *Concrete at High Temperatures: Material Properties and Mathematical Models*. Addison-Wesley, London

Bažant Z P, Prasannan S 1989 Solidification theory for concrete creep: Parts I and II. *J. Eng. Mech.* **115** (8), 1691–703 and 1704–25

RILEM Committee TC 69, Bažant Z P (ed.) 1998 *Mathematical Modeling of Creep and Shrinkage of Concrete*. Wiley, Chichester, UK, Chaps. 2–3, pp. 57–215

RILEM Committee TC 107, Acker P 1998 Measurement of time-dependent strains of concrete. *Mater. Struct.* **31** (212), 507–512

Wittmann F H 1982 Creep and shrinkage mechanisms. In: Bažant Z P, Wittmann F H (eds.) *Creep and Shrinkage of Concrete Structures*, Wiley, London, pp. 129–161

Wittmann F H, Roelfstra P E 1980 Total deformation of loaded drying concrete. *Cement Concrete Res.* **10**, 211–24

Xi Y, Bažant Z P, Jennings H M 1994 Moisture diffusion in cementitious materials. Parts I & II. *Adv. Cement Based Mater.* **1**, 248–257, 258–66

Z. P. Bažant
Northwestern University, Evanston, Illinois, USA

Creep of Particle Reinforced Materials

Two types of particle-strengthened materials are used in high-temperature applications:

- Precipitation and dispersion-strengthened materials with typical particle dimensions below about 100 nm. As interparticle spacing is small, strengthening occurs by direct interaction with dislocations. Particles causing this type of strengthening are referred to as small throughout this paper.
- Composite materials containing discontinuous reinforcements with typical dimensions above about 1 μm. Since the back stress exerted on moving dislocations is insignificant as a result of the large interparticle spacing, strengthening does not occur by direct interaction with dislocations. Instead, load transfer from the softer matrix onto the harder particles is the underlying mechanism. This particle class will be referred to as large in the following.

In addition to direct strengthening, both particle types may also have an indirect effect as they modify the microstructure (e.g., grain size and dislocation substructure) relative to the unreinforced matrix. The goal of this article is to discuss and compare the underlying strengthening mechanisms, concentrating on direct strengthening by small and large particles in Sects. 1 and 3 while discussing indirect effects in Sect. 2. Finally, interaction of small and large particles in so-called dual scale particle-strengthened materials is addressed in Sect. 4.

1. Creep Strengthening by Small Particles

1.1 The Climb Process

The most obvious cause for particle strengthening at high temperatures is the need for dislocations to circumvent particles by climb. It is therefore not surprising that early theories concentrated on this mechanism. Historically, there have been two schools of thought regarding climb geometries: in the so-called local climb model developed by Brown and Ham (1971), it is assumed that the climbing dislocation segment profiles the dispersoid while the dislocation between the dispersoids remains in its glide plane (Fig. 1(a)). In contrast, general climb models (Lagneborg and Bergmann 1976, McLean 1985, Rösler and Arzt 1988) assume that the dislocation is allowed to unravel, i.e., to reduce, under the action of its line tension, the high curvature at the point where it meets the dispersoid (Fig. 1(b)). A particular attraction of the local climb concept was that it offers a simple explanation for the extreme stress sensitivity of the creep rate found in dispersion-strengthened materials: as a significant length of new dislocation line has to be created in the course of dislocation climb, a relatively large threshold stress σ_{th} is predicted which depends on particle geometry and is roughly given by $\sigma_{\mathrm{th}} \approx 0.4 \ldots 0.7\sigma_0$. Nevertheless it is known today that this coincidence is fortuitous and that the special creep behavior of dispersion-strengthened materials has to be related to the detachment barrier caused by an attractive particle-dislocation interaction (see below). Furthermore, it is apparent that the sharp bend in the dislocation line (see point C, Fig. 1(a)), instrumental in producing a significant threshold stress, is extremely unstable and can be rapidly relaxed by diffusion, leading to more general climb as schematically depicted in Fig. 1(b).

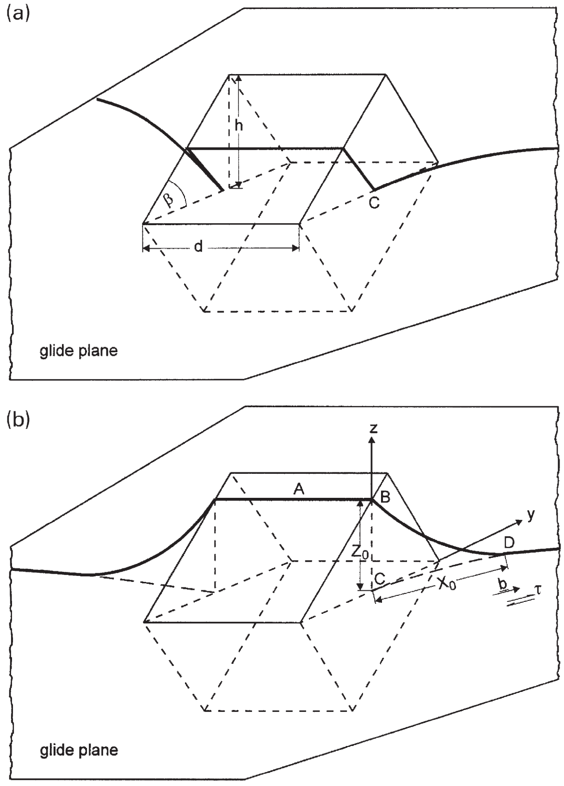


Figure 1
Schematic profile of a dislocation climbing over a cube shaped particle according to the local climb assumption (a) with a sharp bend at point C and the general climb assumption (b) where the high curvature at C is relaxed by climb of the dislocation segment adjacent to the particle. Geometrical details of a climbing dislocation identifying particle height *h*, particle diameter *d*, ramp angle *β*, and the dislocation shape parameters x_0, z_0 are also shown.

Accepting the general climb concept as relevant, estimates of creep retardation by dislocation climb can be made. Noting that the length l of the climbing dislocation segment BD (Fig. 1(b)) is increasing in the course of climb, a back stress τ_{back} results that can be calculated provided the shape of the climbing dislocation segment is known as a function of applied stress, climb position z_0, and particle geometry:

$$\tau_{\mathrm{back}} = \frac{Gb}{2\lambda}\left(\frac{\mathrm{d}l}{\mathrm{d}y}\right) \tag{1}$$

($\mathrm{d}l/\mathrm{d}y$: increase in dislocation line length $\mathrm{d}l$ per advance $\mathrm{d}y$ of dislocation segment BD in glide direction y). Modeling the particles as noninteracting (no repulsive force due to misfit stresses or elastic constant misfit), Lagneborg finds that the maximum back stress τ_{back}^{max}, which is attained when the dislocation has climbed to the top of the particle, is proportional to the applied stress

$$\tau_{\mathrm{back}}^{\max} \approx 0.7\tau \tag{2}$$

but practically independent of particle size and interparticle spacing. Analyzing the climb geometry by requiring equilibrium conditions, i.e., constant chemical potential for vacancies along the climbing segment ABD (Fig. 2), Rösler and Arzt (1988) derived an implicit expression for equilibrium shape $x_0(z_0)$

$$\frac{x_0}{d} = \left[\sqrt{1-(\tau/\tau_0)^2}\left(\frac{x_0}{z_0}+\frac{z_0}{x_0}\right) \times \left(\frac{\tau}{\tau_0}\frac{1}{\tan\beta} - 1 + \left(1-\frac{z_0}{x_0}\right)9/4\right)\right] - 1 \tag{3}$$

from which the line length differential $\mathrm{d}l/\mathrm{d}y$ can be easily calculated. Here, d and β are diameter and ramp angle of a cuboidal particle. The results, showing the normalized back stress $\tau_{\mathrm{back}}/\tau$ as a function of normalized dislocation position z_0/d, are given in Fig. 2. Although a significantly lower back stress ($\tau_{\mathrm{back}}^{\max} \approx 1/3\tau$) is predicted, the qualitative findings are in good agreement with Lagneborg's earlier work (1976)

- the maximum back stress $\tau_{\mathrm{back}}^{\max}$ is attained at the particle top; and

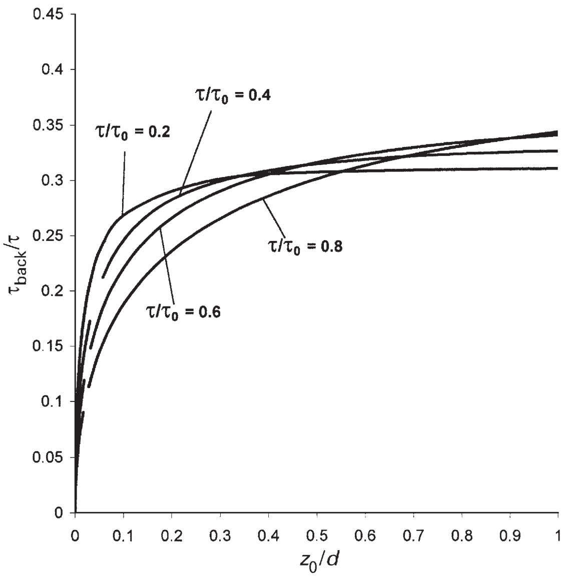


Figure 2
Calculated normalized back stress τ_{back}/τ as a function of normalized climb position z_0/d for different normalized applied stress levels τ/τ_0 (ramp angle $\beta = 45°$).

- τ_{back}/τ is independent of particle diameter d and weakly dependent on Orowan stress, i.e., interparticle spacing 2λ.

The latter point may be best understood by the following example: imagine a situation where the particle volume fraction is doubled at constant particle size. As τ/τ_o is reduced accordingly at given applied stress τ, the unraveling distance x_0 increases so that less extra line length is needed for climb. In consequence, force in forward direction $\vec{F} = \tau\lambda b$ and retarding force, associated with the need to create additional dislocation line length, decrease at the same time so that the overall effect of interparticle spacing on τ_{back}/τ is small. In the context of the back stress concept, this means that creep retardation by particles is largely independent of particle parameters. However, it is important to note that a number of limitations apply. First, it is apparent that τ_{back} cannot exceed the particle cutting or Orowan stress so that the onset of measurable creep retardation is dependent on particle parameters, while the magnitude of strengthening by climb at sufficiently low stresses is not.

Furthermore, it is important to note that general climb models are based on the assumption of equilibrium dislocation shape, i.e., minimum dislocation energy, which requires that vacancies are provided rapidly enough. This prerequisite breaks down when particles are too large and/or deformation rates too fast so that back stress levels calculated under the above assumptions are no longer meaningful. In such a situation it may be more appropriate to consider climb past particles as rate controlling and derive the creep kinetics via

$$\dot{\gamma} = \frac{\rho b 2\lambda}{t_c} \tag{4}$$

where t_c is the time required to surmount a particle by climb and 2λ is the mean free dislocation path, which is equivalent to the interparticle spacing. Following the analysis by Rösler and Arzt (1988) leads to

$$\frac{1}{t_c} = \frac{Gb^4}{d^3} C_i A \left(\frac{\tau - \tau_{\text{back}}}{\tau_o} \right)^n \tag{5}$$

where stress exponent n and preexponential factor A are weakly dependent on particle height h and ramp angle β as

$$n = 3.5\beta^{0.3}(h/d)^{0.2} \tag{6}$$

and

$$A = 60 \times 10^{-1.9\beta}(d/h)^{1.6} \tag{7}$$

Note that a stress exponent $n \approx 3$ results for typical particle parameters. If a stress exponent of two, accounting for a quadratic stress dependence of dislocation density, is added, then n lies in a range typical of the stress dependence of the creep rate in precipitation-strengthened alloys at low stresses. As noted by Lagneborg, the frequently observed break in the strain rate curves of precipitation-strengthened materials to larger n-values at high stresses can be associated with the transition from climb to particle cutting or Orowan bowing irrespective of the applied climb model.

C_i in Eqn. (5) is a kinetic constant given by $C_v = 2\pi D_v d/kTb$ and $C_p = a_p D_p/kT\lambda b$ for volume and pipe diffusion control. In the original analysis, τ_{back} is associated with a threshold stress occurring because a dislocation threads over and under neighboring particles, limiting the unraveling distance x_0. This effect, however, is insignificant and can be neglected in most situations. Hence, it may be of more practical significance to associate τ_{back} to other matrix strengthening mechanisms, such as solid solution hardening.

Inserting for C_i and τ_0 and taking $n = 3$, we find that the dislocation velocity $v = 2\lambda/t_c$ is proportional to $(d/f)^2$ and $(1/f)^{3/2}$ for volume and pipe diffusion, respectively (where f is particle volume fraction). This demonstrates the strong influence of particle parameters on creep rate once climb past particles is rate controlling. In conclusion, it is noted that there are two limiting cases. For small particle dimensions and volume fraction, particle strengthening may be best described by a back stress of the order $\tau_{\text{back}} \approx 0.3 \ldots 0.7$ independent of particle parameters.

Once vacancy supply for dislocation climb past particles becomes rate limiting, particle parameters have a pronounced effect and the creep rate may be best modeled using Eqn. (4) with (5).

1.2 The Detachment Process

The elastic constant of thermally stable ceramic particles such as carbides and oxides is usually larger than that of the metallic matrix. For this reason, one would expect a repulsive interaction between dislocations and particles. However, it was shown by Srolovitz *et al.* (1983) using computer simulation that an attractive interaction results at high temperature provided the particle/matrix interface is incoherent, so that part of the dislocation stress field can be relaxed at the particle by short-range diffusion, whereby the incoherent interface serves as source and sink for vacancies. The theoretical prediction is confirmed by TEM studies (Nardone and Tien 1983, Schröder and Arzt 1985), which show that dislocations are captured at the detachment side of dispersoid particles after climb has been completed. Hence, one has to distinguish between noninteracting particles where creep retardation results merely from climb and particles exerting an attractive interaction on dislocations so that detachment also has to be considered. Since stress relaxation at interfaces requires absorption and emission of vacancies, the interaction strength between particles and dislocations depends on the degree of incoherency whereby coherent particles stand for the one extreme of noninteracting particles. The other extreme is represented by stable voids allowing for maximum stress relaxation.

The interaction strength can be described by the so-called relaxation factor κ defined as

$$\kappa = \frac{T_\mathrm{p}}{T_\mathrm{m}} \tag{8}$$

where T_p and T_m are dislocation line tension at the particle and in the matrix, respectively. For $\kappa = 1$, no attraction between particle and dislocation exists because line energy is not relaxed. For $\kappa < 1$, an attractive interaction results which becomes stronger with diminishing κ. Analyzing the energetics of dislocation detachment, a simple relation between maximum back stress τ_d and κ was found by Arzt and Wilkinson (1986)

$$\tau_\mathrm{d} = \tau_0\sqrt{1-\kappa^2} \tag{9}$$

whereby τ_d is always obtained at the point of dislocation detachment from the particle.

It is important to note that τ_d is not a true threshold stress for creep. As a result of the stress-distance profile exhibiting a sharp stress peak at the point of detachment, the extra activation energy needed for detachment to occur well below τ_d is small. Thus, dislocation detachment has to be treated as a thermally activated process. Taking this into account, Rösler and Arzt (1990) have derived a constitutive creep law

$$\dot{\varepsilon}_\mathrm{d} = \dot{\varepsilon}_0 \exp\left(-\frac{Gb^2 r[(1-\kappa)(1-\sigma/\sigma_\mathrm{d})]^{3/2}}{kT}\right) \tag{10}$$

whereby $\dot{\varepsilon}_\mathrm{d}$ is the creep rate for detachment-controlled creep and the numerator of the exponential term represents the activation energy E_a for dislocation detachment. The reference strain rate is given by $\dot{\varepsilon}_0 = 6D_\mathrm{v}\lambda\rho/b$ (ρ = dislocation density, r = dispersoid radius). As any equation treating a creep deformation process, it can apply strictly only if (i) contributions of other strengthening mechanisms can be neglected and (ii) other deformation mechanisms such as diffusional creep are not operating. It is important to bear this in mind and carefully examine the underlying creep mechanisms when attempting quantitatively to fit experimental creep data of dispersion-strengthened materials to the above equation. However, in any case a number of important conclusions pertaining to alloy development can be made from Eqn. (10):

- since the activation energy E_a for dislocation detachment passes through a maximum as a function of particle radius r (note that σ_d is dependent on r), an optimum particle size exists which contrasts predictions of classical threshold stress concepts. Physically, this results because (i) the glide distance between the equilibrium position of the dislocation and their point of detachment is proportional to the particle dimensions and, thus, E_a is proportional to r at given normalized stress σ/σ_d and (ii) larger particles are associated with a lower Orowan stress, i.e., athermal detachment stress σ_d, at given volume fraction. Analytically, the optimum particle size is given by $(r/b)_\mathrm{opt} = (5/3(1-\kappa))^{3/2}(\ln(\dot{\varepsilon}_0/\dot{\varepsilon})kT/Gb^3)$. Depending on loading conditions and κ-value, $(r/b)_\mathrm{opt} = 10 \ldots 100$ is typically found.
- Low values of k are essential for superior high-temperature performance, because the influence of the κ-factor in Eqn. (10) is twofold (note that σ_d is dependent on κ).
- As the normalized stress exponent n/r of the creep curve is a good measure for κ, optimum dispersoid/matrix combinations can be identified from creep curve analysis: the steeper the creep curve at fixed particle size and test temperature, the stronger the attractive force between particles and dislocations.

Furthermore, it is important to note that experimental observations of temperature dependent threshold stresses and activation energies exceeding values for self diffusion (Kucharova *et al.* 1988, Rösler *et al.* 1992) are readily explained by taking thermal activation into account.

Dispersion strengthening is usually dominant at low strain rates and very high temperatures due to the very high stress sensitivity of the creep rate. However, at faster deformation rates other hardening mechanisms have to be considered as well. One possibility is to assume a sequence of glide steps whereby a dislocation by-passes a dispersoid in step one at creep rate $\dot{\varepsilon}_1$, defined by Eqn. (10), and glides a distance 2λ between particles in a second step at creep rate $\dot{\varepsilon}_2$, given by the power law equation of the matrix. The overall behavior is then described as $1/\dot{\varepsilon} = 1/\dot{\varepsilon}_1 + 1/\dot{\varepsilon}_2$. Another approach is to associate a back stress $\sigma_{\text{back}i}$ with each individual strengthening mechanism, so that the effective stress to be inserted into the power law equation of the unreinforced matrix is given by

$$\sigma_{\text{eff}} = \sigma - \sum_i \sigma_{\text{back,i}} \tag{11}$$

Following a suggestion by Reppich *et al.* (1999), we can rewrite Eqn. (10) resulting in an expression for the back stress caused by thermally activated dislocation detachment from dispersoids

$$\frac{\sigma_{\text{back}}(\dot{\varepsilon},\ T)}{\sigma_{\text{d}}} = 1 - \frac{1}{1-\kappa}\left[\frac{kT}{Gb^2 r}\ln\left(\frac{\dot{\varepsilon}_0}{\dot{\varepsilon}}\right)\right]^{2/3} \tag{12}$$

Note that σ_{back} decreases the higher the temperature and the lower the strain rate which, once again, demonstrates the absence of a threshold below which creep deformation vanishes.

In Fig. 3, both procedures are compared for material data given in the figure caption. As expected, the back stress concept leads to somewhat higher strength estimates because it assumes a long range back stress associated with dispersoids, while the first approach assumes free dislocation glide between obstacles. It also predicts a more gradual transition from matrix like creep (low stress exponent) to detachment-controlled deformation (high stress exponent) as often observed experimentally. Because arguments can be found for both concepts, it appears to be a matter of taste which one to use. Reppich *et al.* (1999) have demonstrated that dispersion-strengthened materials can be successfully described using the back stress concept and considering thermally activated dislocation detachment.

2. *Effect on Microstructure*

In addition to direct strengthening, small and large particles can also indirectly influence creep performance because they affect matrix microstructure, in particular dislocation density and (sub-)grain size (Gonzalez-Doncel and Sherby 1993, Heilmaier *et al.* 1996, Lin and Sherby 1981, Reppich *et al.* 1999). A popular approach to correct for substructure stabilization by particles is to apply Sherby's constant structure steady-state creep equation (Sherby *et al.* 1977), which was developed for cases where subgrain size does not vary with stress. However, detailed measurements often show a clear stress dependence (Reppich *et al.* 1999). Hence, it appears to be more accurate (but also more tedious) to use stress dependent laws for substructure evolution, which have to be calibrated to experimental data and, by that, quantify the back stress exerted on moving dislocations.

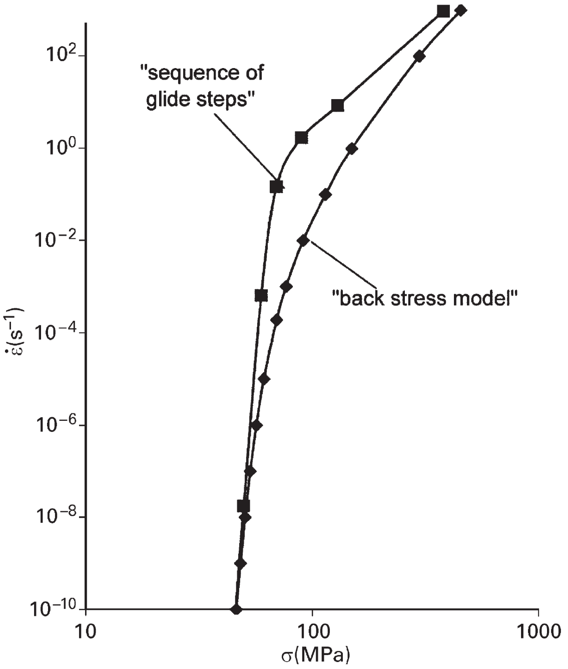


Figure 3
Calculated strain rate $\dot{\varepsilon}$ as a function of applied stress for dispersion-strengthened aluminum containing 5% dispersoids with radius $r = 15\,\text{nm}$ and relaxation factor $\kappa = 0.9$ (all material parameters according to (Frost and Ashby 1982), dislocation density ρ: $1 \times 10^{13}\,\text{m}^{-2}$). The curve labeled back stress model results when an effective stress according to Eqns. (11) and (12) is inserted in the power law creep equation of the matrix. For the curve labeled sequence of glide steps it is assumed that the dislocation bypasses a dispersoid at creep rate given by Eqn. (10) and glides between particles at creep rate given by the power law for matrix creep.

Microstructural stabilization is not necessarily beneficial, especially in the case of dispersion-strengthened materials operating at extreme temperatures. Often grain sizes in the range of a few micrometers are stabilized (Kucharova *et al.* 1988) which promote diffusional creep. This effect is clearly seen by comparing fine- and coarse-grained dispersion-strengthened materials (Gregory *et al.* 1985, Lin and Sherby 1981, Rösler *et al.* 1992). However, a striking observation is that (i) deformation rates are grossly overestimated using Coble– and Nabarro–Herring creep equations and (ii) stress exponents $n \gg 1$, but smaller than in the coarse grained material, are observed. Some authors have suggested a coupling between grain boundary sliding and dislocation creep as a possible cause. However, it has been demonstrated by finite element calculations that, except in a narrow transition region, the stress exponent is still determined by dislocation creep when such a coupling takes place. A second model capable of explaining anomalously high stress exponents is that of interface reaction-controlled diffusional creep (Arzt *et al.* 1983). It assumes that vacancy generation and annihilation require the motion of grain boundary dislocations, which have to bypass dispersoids in the grain boundary. Assuming that this bypass process is, as in the bulk material, controlled by thermally activated dislocation detachment, a rate equation can be derived (Rösler *et al.* 1992)

$$\dot{\varepsilon}_{\mathrm{diff}} = \frac{4\lambda\rho_{\mathrm{gb}}^{\mathrm{e}}D_{\mathrm{v}}}{d_{\mathrm{g}}b_{\mathrm{gb}}} \cdot \exp\left(-\frac{Gb_{\mathrm{gb}}^{2}r[(1-\kappa)(1-\sigma/\sigma_{\mathrm{d}})]^{3/2}}{kT}\right) \tag{13}$$

where ρ_{gb}^{e} is the density of grain boundary edge dislocations and $b_{\mathrm{gb}} \approx 0.3 \ldots 0.5b$ is a typical estimate of the Burgers vector of a grain boundary dislocation. As the stress exponent n scales with the square of the Burgers vector for both detachment-controlled diffusional creep and detachment-controlled dislocation creep (Eqns. (10) and (13)), the stress exponent should be reduced by $(b_{\mathrm{gb}}/b)^2 \approx 0.1 \ldots 0.25$ when the rate controlling mechanism changes accordingly. This prediction matches well with experimental observations made for dispersion-strengthened aluminum and superalloys, although direct evidence for interface reaction-controlled diffusion creep is still missing.

3. Creep Strengthening by Large Particles

In Sect. 1, retardation of dislocation motion by small particles was discussed. In all cases, the Orowan stress has to be a significant fraction of the applied stress for noticeable strengthening to occur. This condition is not normally met for large reinforcements in the micrometer to millimeter size range because interparticle spacing is too large to exert an appreciable back stress on moving dislocations. Nevertheless, direct strengthening is observed, in addition to indirect strengthening as discussed above, due to load transfer from the soft matrix onto the hard reinforcements (Christman *et al.* 1998, Kelly and Street 1972, Dragone and Nix 1990). Moreover, hydrostatic stresses develop in the matrix as a consequence of constrained plastic flow, which increase the load carrying capacity of the composite as the deviatoric stresses in the matrix are reduced. Consequently, for matrix materials exhibiting power law creep, reinforcement strengthening by large particles can be accounted for using a modified power law for steady-state creep

$$\dot{\varepsilon} = A\frac{D_{\mathrm{v}}Gb}{kT}\left(\frac{\sigma}{\Lambda G}\right)^{n} \tag{14}$$

whereby $\Lambda > 1$ is the strengthening coefficient dependent on reinforcement volume fraction and aspect ratio, as well as stress exponent n. Λ can be readily interpreted as the factor by which the effective flow stress is reduced due to load transfer and build-up of hydrostatic stresses in the matrix. Trends in Λ have been calculated by Bao *et al.* (1991) using finite element analysis for aligned, axisymmetric reinforcements and approximated by analytical equations for stress exponent $n \to \infty$ (Rösler *et al.* 1991), giving

$$\Lambda \approx 1 + 2(2 + l/R)f_{\mathrm{r}}^{3/2} \tag{15}$$

for fibers with length $2l$, diameter $2R$ and reinforcement volume fraction f_{r}. It should be noted that particle distribution plays an important role and thus, calculations assuming idealized arrangements capture trends, but not the precise situation in a particular material. Nevertheless, Eqn. (15) demonstrates that high aspect ratios and volume fractions are key in attaining significant strengthening. Considering SiC-whisker-strengthened aluminum with $l/R = 5$ and $f_{\mathrm{r}} = 13\%$ for example, $\Lambda = 1.56$ results from Eqn. (15) which is in good agreement with experimental observations. Since Λ is independent of applied stress, theory predicts identical stress exponents as in the unreinforced matrix material. Of course, the real situation may be more complicated due to indirect strengthening effects and damage accumulation, which may be initiated in a very early state of the total lifetime.

Interestingly, absolute particle dimensions do not come into play in the above analysis, although it is generally accepted that they must be large compared with the microstructural scale for continuum mechanics to apply. Hence, the question arises to which extent particles can be reduced in size and still contribute to strengthening by load transfer. This is of particular practical significance because many materials, such as particulate-reinforced aluminum alloys, contain fairly small reinforcements in the micrometer range (Gonzalez-Doncel and Sherby 1993). It has been argued that, at high temperatures, a lower limit for acceptable particle size results from the onset of diffusional creep of the volume element associated with the reinforcement. Regardless of particle strength, the volume element can be deformed by vacancy flux, much as a single grain included in an otherwise single crystalline material. As it is sufficient to move matrix atoms around the particle, the mechanism does not require shape change of the particle by diffusion of reinforcement atoms. Only when the creep rate by diffusion is negligible compared to the matrix creep rate, can the full strengthening contribution due to load transfer be expected. In contrast, if diffusional creep is infinitely fast, the volume element acts like a void and weakening occurs. By equating both deformation rates, a critical dimensionless particle parameter χ can be identified below which reinforcement strengthening is eliminated at high temperatures (Rösler *et al.* 1991)

$$\chi = \frac{Rl^2}{\Omega} = \frac{4(\delta D_b + D_v l/\pi)}{kT}\frac{\sigma}{\dot{\varepsilon}} \qquad (16)$$

(using an effective diffusion coefficient).

Here, $\dot{\varepsilon}$ is the creep rate of the unreinforced matrix material at stress σ. The predictions of the analytical Eqn. (16) have been precisely reproduced by finite element calculations for a number of particle aspect ratios and volume fractions (Nimmagadda and Sofronis 1996).

Taking a PM 6092 aluminum alloy operating at 350 °C and $\dot{\varepsilon} = 10^{-8}\,\mathrm{s}^{-1}$ as an example, a strength of about 10 MPa results. Using diffusion data from Frost and Ashby (1982) and assuming that interface and grain boundary diffusion coefficients are roughly identical, $r \approx 40\,\mu\mathrm{m}$ is obtained from Eqn. (16) as a critical dimension for equiaxed particles. In the light of this result it seems surprising that aluminum composites containing SiC, TiC, or Al_2O_3 particles with dimensions typically in the range of a few micrometer exhibit superior strength at temperatures as high as 500 °C. However, one has to remember that hardening may also stem from indirect effects and, in fact, there is no common opinion in the literature whether direct or indirect strengthening is responsible for the observed behavior. Since it is hard to imagine a mechanism preventing diffusional creep at the location of the particles, it is reasonable to conclude that strengthening by load transfer cannot account for strengthening at extreme temperatures, unless significantly larger particles are used. In particular, the high stress sensitivity of the creep rate observed for composites processed by powder metallurgy should not be seen as argument for direct strengthening because it is very likely that this effect is caused by dispersoids introduced during processing (Park *et al.* 1990).

Focusing on direct strengthening for the moment, it is interesting to compare small and large particles. Small particles causing attractive interaction with dislocations are best for applications at very low normalized strain rates $\dot{\varepsilon}/D_v$ because of the very high stress exponent associated with dislocation detachment. No other strengthening mechanism produces similar effects. In the absence of an attractive interaction, the situation is less obvious. According to Lagneborg's results (Lagneborg and Bergmann 1976) and the above analysis, the effect of small particles on creep may be approximated assuming a back stress of about 0.5σ for the limiting case of low volume fractions and stresses below the particle cutting or Orowan stress (hereby, the effect of climb kinetics is neglected, see Sect. 1.1). Identical strengthening by large reinforcements would require $\Lambda = 2$ (see Eqn. (14)). It is easily seen from Eqn. (15) that this requires high aspect ratio reinforcements because unrealistically high volume fractions of more than 30% would be needed using equiaxed particles. Although strengthening by large equiaxed particles appears to be comparably unattractive from this point of view, there are also clear advantages: (i) strengthening is effective over the entire stress range, whereas the back stress exerted by small particles is insignificant for $\sigma \gg \sigma_o$; (ii) thermal stability of ceramic reinforcements far exceeds the stability of coherent precipitates; and (iii) large reinforcement volume fractions improve elastic stiffness and wear resistance in addition to high-temperature performance. Hence, it is primarily a matter of application temperature and overall performance requirements whether to rely on precipitation or reinforcement strengthening.

4. *Dual Scale Particle-strengthened Materials*

It has recently been suggested (Rösler and Bäker 2000) that dual scale particle-strengthened materials, containing small and large particles in appropriate combination, may outperform dispersion-strengthened materials, containing only small particles, and composites, containing only large particles. The rationale is as follows: considering a dispersion-strengthened matrix to which large reinforcements are added, it is clear from the above discussion that the effective stress causing dislocation motion and, in particular, dislocation detachment from dispersoid particles is reduced by the factor Λ. Hence, the creep law for a material containing dispersoid particles and large reinforcements under detachment-controlled creep can be simply expressed as

$$\dot{\varepsilon}_{\mathrm{d,c}} = \dot{\varepsilon}_0 \exp\left(-\frac{Gb^2 r[(1-\kappa)(1-\sigma/\Lambda\sigma_{\mathrm{d}})]^{3/2}}{kT}\right) \qquad (17)$$

where $\dot{\varepsilon}_{d,c}$ is the creep rate of the dual scale particle-strengthened material. For a given dispersion-strengthened matrix, particle radius r and relaxation factor κ are fixed, so that the critical factor determining creep strength is $\Lambda\sigma_d = \Lambda\sigma_0\sqrt{(1-\kappa^2)} \propto (1 + 2(2 + l/R)f_r^{3/2})\sqrt{f_d}$ (f_d: dispersoid volume fraction, f_r: reinforcement volume fraction). Interestingly, $(1 + 2(2 + l/R)f_r^{3/2})\sqrt{f_d}$ goes through a maximum at fixed total particle volume fraction $f_d + f_r$ under certain conditions. For example, setting $f_d + f_r = 0.2$ a numerical value of 0.45 is obtained if the material is entirely strengthened by dispersoids. However, if 15%

reinforcements are combined with 5% dispersoids, 0.54 and 0.30 are obtained as numerical values for $l/R = 10$ and $l/R = 1$, respectively. This means that a strength elevation of about 20% is predicted by dual scale particle strengthening for the example of high aspect ratio particles. In contrast, equiaxed particles are of no benefit. At this point, the prediction is merely theoretical and it remains to be seen whether the anticipated effect can be utilized in practice.

5. Conclusions

The effect of small and large particles on creep performance by direct and indirect strengthening mechanisms has been discussed. Small particles cause direct hardening by interaction with dislocations. Hereby, two barriers have to be considered: dislocation climb and dislocation detachment. Creep retardation by climb is caused because additional dislocation line length must be created in order to surmount particles. In the limiting case of low particle volume fractions, a back stress of the order 0.3 … 0.7σ results. This is the only direct strengthening mechanism of coherent particles. In contrast, incoherent dispersoids cause an additional detachment barrier because they exert an attractive force on moving dislocations. The creep curve is characterized by a remarkably high stress sensitivity, which should, however, not be mistaken as sign for the existence of a true threshold stress.

Large reinforcements do not cause strengthening by direct interaction with dislocations. Instead, load transfer from the softer matrix onto the harder particles is the underlying mechanism. The creep behavior can be readily described by the power law for steady-state creep of the unreinforced matrix, but the applied stress σ must be replaced by an effective stress σ/Λ. Hereby $\Lambda > 1$ is the strengthening coefficient which reflects the fact that the effective flow stress is reduced due to load transfer and built-up of hydrostatic stresses in the matrix. To make full use of reinforcement strengthening at high temperatures, it is important to select sufficiently large particle dimensions in order to avoid stress relaxation at the particle by diffusional creep.

Bibliography

Arzt E, Ashby M F, Verall R A 1983 Interface controlled diffusional creep. *Acta Metall.* **31**, 1977–89

Arzt E, Wilkinson D S 1986 Threshold stresses for dislocation climb over hard particles: the effect of an attractive interaction. *Acta Metall.* **34**, 1893–8

Bao G, Hutchinson J W, McMeeking R M 1991 Particle reinforcement of ductile matrices against plastic flow and creep. *Acta Metall. Mater.* **39**, 1871–82

Brown L M, Ham R K 1971 Dislocation-particle interactions. In: Kelly A, Nicholson R B (eds.) *Strengthening Methods in Crystals.* Elsevier, Amsterdam, pp. 9–135

Christman T, Needleman A, Suresh S 1998 An experimental and numerical study of deformation in metal-ceramic composites. *Acta Metall.* **37**, 3029–50

Crossmann F W, Ashby M F 1975 The non-uniform flow of polycrystals by grain-boundary sliding accommodated by power-law creep. *Acta Metall.* **23**, 425–40

Dragone T L, Nix W D 1990 Geometric factors affecting the internal stress distribution and high temperature creep rate of discontinuous fiber reinforced metals. *Acta Metall.* **38**, 1941–53

Frost H J, Ashby M F 1982 *Deformation Mechanism Maps.* Pergamon, Oxford

Gonzalez-Doncel G, Sherby O D 1993 High temperature creep behavior of metal matrix aluminium-SiC composites. *Acta Metall. Mater.* **41**, 2797–805

Gregory J K, Gibeling J C, Nix W D 1985 High temperature deformation of ultra-fine-grained oxide dispersion strengthened alloys. *Metall. Trans.* **16A**, 777–87

Heilmaier M, Wunder J, Böhm U, Reppich B 1996 Application of the HAI-model to the creep behaviour of particle hardened nickel-base superalloys. *Computational Materials Science* **7**, 159–66

Kelly A, Street K N 1972 Creep of discontinuous fibre composites, II. Theory for the steady-state. *Proc. R. Soc. Lond.* **328A**, 283–93

Kucharova K, Orlova A, Oikawa H, Čadek J 1988 Creep in an aluminium alloy strengthened by Al_4C_3 particles. *Mater. Sci. Eng.* **102A**, 201–9

Lagneborg R, Bergmann B 1976 The stress/creep rate behaviour of precipitation-hardened alloys. *Met. Sci.* **10**, 20–8

Lin J, Sherby O D 1981 Creep of oxide dispersion strengthened materials (with special reference to T-D nichrome). *Res. Mechanica.* **2**, 251–93

McLean M 1985 On the theshold stress for dislocation creep in particle strengthened alloys. *Acta Metall.* **33**, 545–56

Nardone V C, Tien J K 1983 Pinning of dislocations on the departure side of strengthening dispersoids. *Scripta. Metall.* **17**, 467–70

Nimmagadda P B R, Sofronis P 1996 Creep strength of fiber and particulate composite materials: the effect of interface slip and diffusion. *Mech. Mater.* **23**, 1–19

Park K T, Lavernia E J, Mohamed F A 1995 High temperature creep of silicon carbide particulate reinforced aluminium. *Acta Metall. Mater. Trans.* **38**, 2149–59

Park K T, Mohamed F A 1995 Creep strengthening in a discontinuous SiC-Al composite. *Metall. Mater. Trans.* **26A**, 3119–29

Reppich B, Heilmaier M, Wunder J, Baumeister H, Huber T 1999 Micromechanical analysis of the creep behaviour of iron-base superalloys. In: Mughrabi H *et al.* (eds.) Microstructure and Mechanical Properties of Metallic High-temperature Materials. Deutsche Forschungsgemeinschaft, Wiley-VCH, pp. 509–30

Rösler J, Arzt E 1988 The kinetics of dislocation climb over hard particles-I. Climb without attractive particle-dislocation interaction. *Acta Metall.* **36**, 1043–51

Rösler J, Arzt E 1990 A new model-based creep equation for dispersion strengthened materials. *Acta Metall.* **38**, 671–83

Rösler J, Bao G, Evans A G 1991 The effects of diffusional relaxation on the creep strength of composites. *Acta Metall. Mater.* **39**, 2733–8

Rösler J, Evans A G 1992 The effect of reinforcement size on the creep strength of intermetallic matrix composites. *Mater. Sci. Eng.* **153A**, 438–43

Rösler J, Joos R, Arzt E 1992 Microstructure and creep properties of dispersion-strengthened aluminium alloys. *Metall. Trans.* **23A**, 1521–39

Rösler J, Bäker M 2000 A theoretical concept for the design of high-temperature materials by dual-scale particle strengthening. *Acta mater.* **48**, 3553–67

Schröder J H, Arzt E 1985 Weak beam studies of dislocation/dispersoid interaction in an ODS superalloy. *Scripta Metall.* **19**, 1129–34

Sherby O D, Klundt R K, Miller A K 1977 Flow stress, subgrain size, and subgrain stability at elevated temperature. *Metall. Trans.* **8A**, 843–50

Srolovitz D J, Petkovic-Luton R A, Luton M J 1983 Edge dislocation-circular inclusion interactions at elevated temperatures. *Acta Metall.* **31**, 2151–9

J. Rösler
Technische Universität Braunschweig, Germany

CVD Monofilaments

Chemical vapor deposition (CVD) is a process in which one material is deposited onto a substrate to produce near-theoretical density and small grain sizes for the deposited material. CVD onto an appropriate fine filamentary substrate produces a thicker filament predominantly composed of materials that otherwise could not be made by drawing or other conventional processes. The fine and dense microstructure of the deposited material ensures maximum strength and Young's modulus. In the early 1960s, interest in the production of fibers that combined a high modulus with low density concentrated efforts on the lightest elements in the Periodic Table. This is because the ultimate strength and Young's modulus of a material are controlled by the outer electrons in the atoms and are not a function of atomic weight. The lightest element that possesses high mechanical properties is the fourth element in the Periodic Table, beryllium, which has a high Young's modulus of over 300 GPa and a density of 1.83. However beryllium is highly toxic, especially as an oxide, and interest centered on the other light elements, i.e., the fifth, sixth, and fourteenth elements, boron, carbon, and silicon.

Boron fibers were the first continuous high performance fibers to be produced. They are made by CVD onto a tungsten substrate or core, and possess a Young's modulus twice that of steel with the density of aluminum. The same technique is also used to produce silicon carbide (SiC) fibers. The latter fibers are produced on both tungsten and carbon cores. Boron and silicon carbide fibers are the only fibers made by CVD and find particular end uses. Their diameters, which are of the order of 20 times those of carbon fibers (i.e., of the order of 100 μm or more), ensure high compressive strength and allow them to be more easily embedded into metal matrices than many other fibers; however, as a result these filaments are also less flexible than other fibers used for reinforcement.

1. Chemical Vapor Deposition

The production of both boron and silicon carbide fibers by CVD takes place in a reactor that consists of a glass tube, which for commercial production is usually 1–2 m in length and most often vertical. Schematic representations of the arrangements for the production of both types of fibers are shown in Fig. 1. Several such reactors can be placed in series so as to allow staged deposition, but in practice this is avoided as the increase in the number of seals increases the risks of contaminants being introduced into the fiber. The wire substrate can run continuously through the reactor by the use of mercury seals at each end. These ensure that the reactive gases are contained in the reactor and also allow the substrate to be heated by electric current over the deposition length. The choice of the substrate is governed by the need to use a material that maintains its strength to around 1100 °C and is an electrical conductor. Heating the substrate, usually with a d.c. current, is a necessary part of the process that determines deposition rate, which must be constant or irregularities in the fiber diameter occur with a subsequent fall in

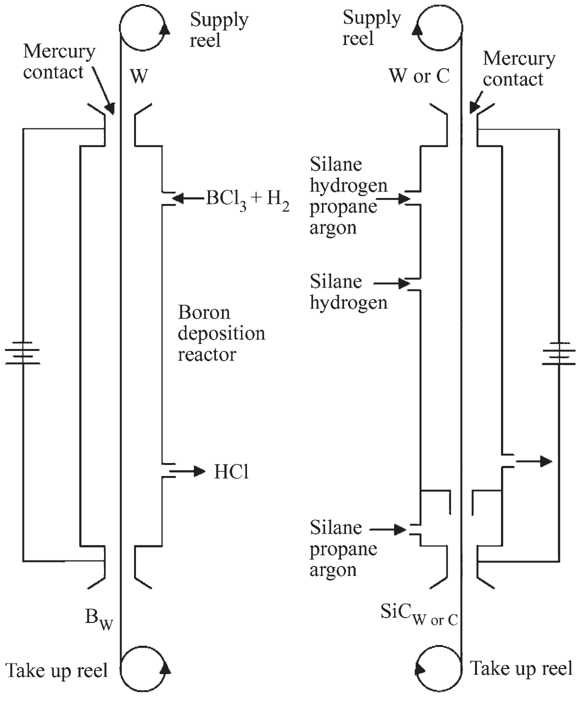


Figure 1
Schematic views of the reactors for producing boron and silicon carbide fibers by CVD.

strength. Several potential substrate materials were investigated including molybdenum and tantalum but tungsten was chosen as it was available in the form of fine wire (produced for electric light filaments) and showed greater compatibility with boron, including having a similar coefficient of thermal expansion. Initial attempts to use carbon cores for the production of boron fibers were not successful due to differences in thermal expansion coefficients, which led to the breakage of the carbon filament. During production, such breakages produce hot spots due to the locally raised electrical resistance in the substrate; in turn, these result in locally enhanced deposition rates and weak spots. This was overcome by the deposition of pyrolytic carbon onto the substrate, which allowed the boron fiber to slide over the carbon fiber. Nevertheless, tungsten filaments are used for the commercial production of boron fibers, whilst both tungsten and carbon filaments are used as substrates for the production of silicon carbide fibers.

2. Boron Fibers

Boron fibers can be made by CVD onto a tungsten core of 12 μm diameter. Several chemical routes are possible but commercially boron trichloride is mixed with hydrogen and boron is deposited according to the following reaction:

$$2BCl_3(g) + 3H_2(g) \rightarrow 2B(s) + 6HCl(g)$$

In commercial production the gas mixture is introduced at the top of the vertical reactor and flows in the drawing direction. The temperature along the reactor passes through a maximum at an early stage of deposition but as Wawner (1988) has shown, it must not exceed 1350 °C as this leads to increased grain sizes and a drop in filament strength. Passage through the reactor is of the order of one or two minutes and results in a fiber with a diameter of 140 μm. During the deposition process the small boron atoms diffuse into the tungsten core to produce complete boridization and the production of WB_4 and W_2B_5. This leads to an increase in diameter to 16 μm, which induces considerable residual stresses, putting the core into compression and the neighboring boron mantle into tension. A final annealing process puts the fiber surface into compression.

Boron fibers can be relatively easily used to reinforce light alloys but it is necessary to protect the fiber surface during contact with the molten metal to avoid chemical reactions occurring with a subsequent fall in strength. For this reason boron fibers have been commercially produced with both silicon carbide and boron carbide (B_4C) coatings.

Boron fibers coated with silicon carbide were originally made with a diameter of 106 μm but were found to split longitudinally due to low transverse strength (Kreider and Prewo 1972). This was avoided by increasing the diameter to 140 μm, but splitting due to localized high stresses when fibers touch during composite manufacture was found to occur even with these latter fibers (Bunsell and Nguyen 1980).

The tungsten wire substrates are produced by conventional wire drawing, which results in their surface being striated. The boron is deposited preferentially onto the highest peaks of the surface of the wire and grows as conical nodules, leading to a nodular surface to the fiber. Figure 2 shows the nodular appearance of a boron fiber. A smoother substrate such as a pitch-based carbon fiber leads to a smoother surface for the deposited fiber. The internal structure of the deposited boron is nanocrystalline with grain sizes of 2–3 nm consisting of α-rhombohedral and tetragonal polymorphs of boron (Bhardwaj and Krawitz 1983). The rate of deposition is strongly dependent on temperature so that any local hot spot will cause enhanced grain growth and surface irregularities, which lead to a fall in strength.

In the absence of such major defects three types of failure initiation processes have been identified. The stress-concentration which is produced by the joining of two neighboring nodules at the fiber surface is the most frequent failure initiation mechanism and surface etching to reduce rugosity results in an increase in strength (Vega-Boggio and Vingsbo 1976a). Two types of defects are associated with the boron mantle–core interface. The growth of neighboring nodules can cause radial defects that are due to the one nodule shadowing the deposition of the boron onto the

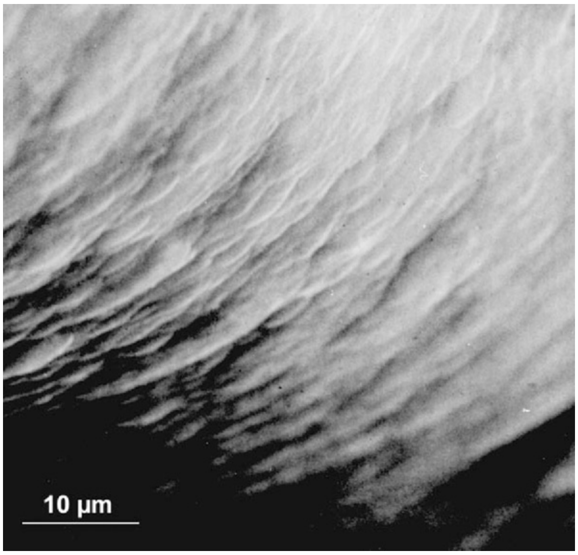


Figure 2
Boron is deposited on the tungsten wire substrate in the form of nodules, which leads to the nodular surface shown.

tungsten wire. Elongated voids can be formed at the interface by preferential deposition on each side of striations on the tungsten wire (Vega-Boggio and Vingsbo 1976b).

Boron fibers are not only strong in tension but also lead to composites that are very strong in compression, probably at least in part due to the large diameter, which inhibits buckling. The strength of the fibers in tension lies between 3.4 and 4.2 GPa with a coefficient of variation of 15% or less, a Young's modulus of 400 GPa, and a specific gravity of 2.54. The Poisson's ratio of boron fibers is 0.13 and their thermal expansion coefficient is $4.9 \times 10^{-6}\,^{\circ}\mathrm{C}^{-1}$ up to 325 °C.

Boron fibers exhibit a linear axial stress–strain relationship at room temperature up to 650 °C. Strength is retained in short-term tests to such high temperatures, but heating in air for periods of hours at temperatures above 300 °C leads to large losses in room temperature strength. This strength loss is attributed to oxidation of the nodular boron surface.

At higher temperatures boron fibers exhibit anelastic behavior, which up to 800 °C can be explained in terms of different strains induced in the boron mantle and the tungsten boride core (Ericksen 1974). A boron fiber formed into a loop and exposed briefly to an open flame, which is then withdrawn, retains its imposed curvature. On reheating, the fiber returns to its original shape, demonstrating that the deformation contains no permanent plastic component. Boron fibers also creep at high temperatures (DiCarlo 1977).

3. *Silicon Carbide Fibers*

Silicon carbide shows high structural stability and strength retention, even at temperatures above 1000 °C. Passive oxidation begins at 1200 °C but in bulk form silicon carbide is used up to 1600 °C. It is therefore an attractive material to have in fiber form for reinforcement in high temperature composite materials. Silicon carbide fibers made by CVD are produced on both tungsten and carbon cores in similar reactors to those use for the production of boron fibers except that multiple injection points are used for the introduction of the reactant gases. Various carbon-containing silanes have been used as reactants. In a typical process, with $MeSiCl_3$ as the reactant, silicon carbide is deposited on the core as follows:

$$MeSiCl_3(g) \rightarrow SiC(s) + 3HCl(g)$$

Silicon carbide fibers produced on a tungsten core of 12 μm diameter show a thin interfacial layer between the silicon carbide mantle and the tungsten core, caused by chemical interaction between the silicon carbide and tungsten, which forms reaction products of α-W_2C and W_5Si_3. Heating above 900 °C leads to growth of this reactive layer and degradation of the properties of the fiber. The carbon fiber cores, of 33 μm diameter, used for the CVD production of silicon carbide fibers, are themselves coated with a 1–2 μm thick layer of pyrolytic carbon, which covers surface defects on the core and produces a more uniform diameter. The fibers on a carbon core show a gradation of composition, which is richer in carbon near the interface and becomes stoichiometric silicon carbide towards the fiber surface. The fibers on a tungsten core have a mantle of stoichiometric silicon carbide. Both types of silicon carbide fibers have smoother surfaces than do boron fibers, which is due to the deposition of small columnar grains rather than conical nodules. The specific gravity of a 100 μm diameter SiC–W fiber is 3.35 whilst that of a 140 μm diameter SiC–C fiber is around 3.2.

Early silicon carbide fibers were found to have a wide scatter in tensile strengths, which was alleviated by coating them with a thin layer of carbon. This had the disadvantage, however, of reducing interfacial bonding, particularly with light alloys. To overcome this, the silicon carbide fibers made on a carbon core and destined to reinforce light alloys, are produced with a surface coating, the composition of which is made to vary from being carbon-rich to silicon carbide at the outer surface (Sattler *et al.* 1992). The fibers that are to be used to reinforce titanium have a protective layer that varies from being rich in carbon, to being rich in silicon, to a composition that is again rich in carbon at the surface (Ning and Pirouz 1991) as shown in Fig. 3. The outer sacrificial layer protects the fiber during contact with the molten and highly reactive titanium during composite manufacture. This type of fiber is called SCS-6, as the coating increases the fiber diameter by 6 μm. The fiber shows no degradation after 5 hours at 900 °C when embedded in a titanium (6Al4V) matrix (Whatley and Wawner

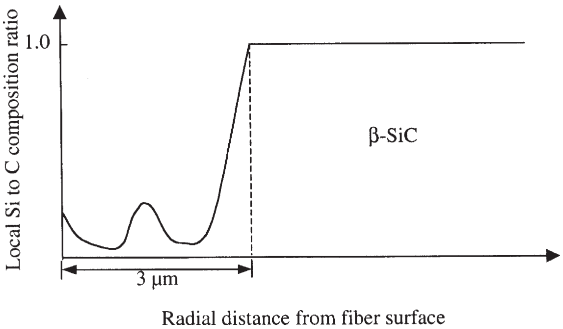


Figure 3
The SiC-6 fiber produced on a carbon core for the reinforcement of titanium has a protective layer is 3 μm thick. This surface layer can be seen to be due to a first coating of carbon, which is then made richer in silicon, followed by a second carbon-rich layer, which plays a sacrificial role during composite manufacture.

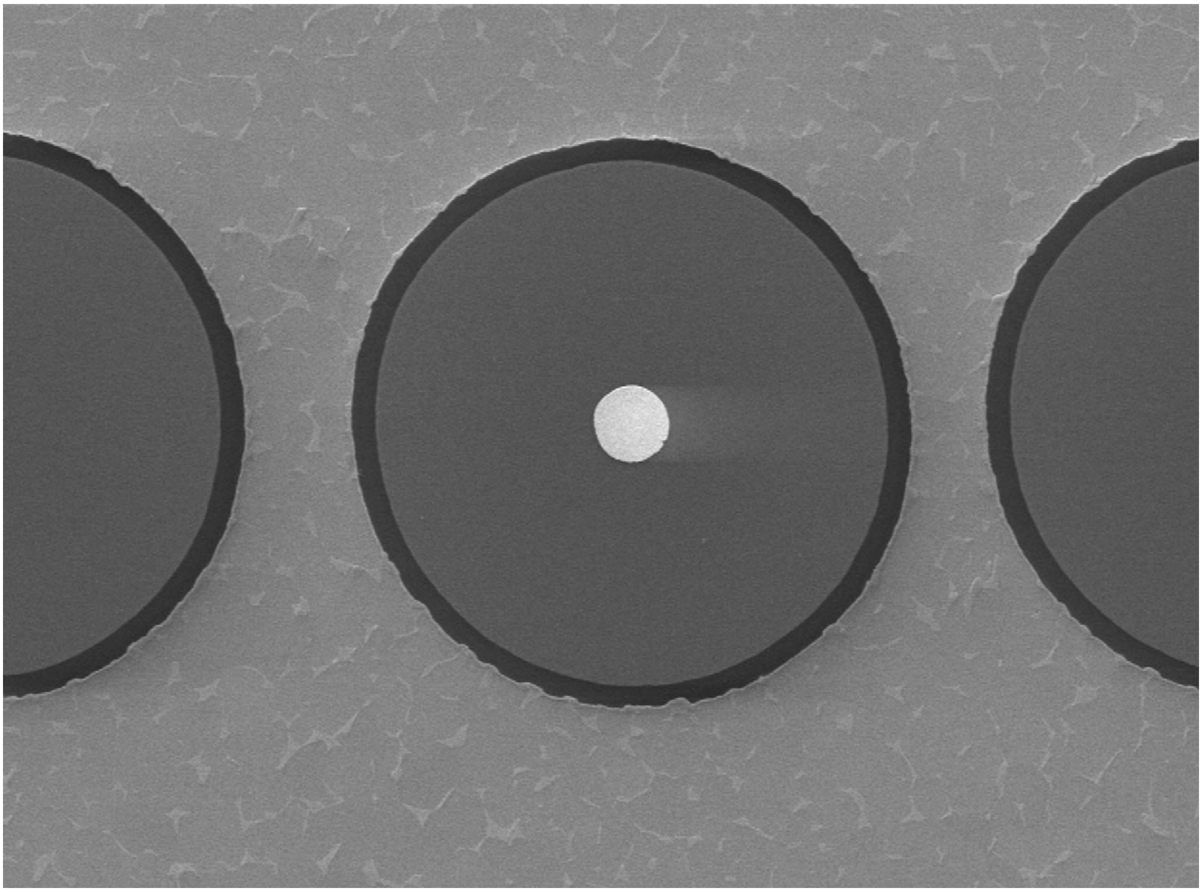


Figure 4
A cross-section of a silicon carbide fiber, having a diameter of 103 μm, produced on a tungsten core and revealing a surface coating of carbon, which has allowed it to be used to reinforce a titanium alloy.

1985). A cross-section of a typical 100 μm diameter silicon carbide fiber on a tungsten core is shown in Fig. 4. This fiber has a surface coating 3–4 μm thick, which varies from being carbon rich to becoming a thin layer of less than 1 μm of TiB_2 at the surface (Shatwell 1999).

The room temperature strengths of the fibers are around 4 GPa. Low failure stresses are due to surface flaws, whereas higher strengths are controlled by defects at the core mantle interface. The strengths of CVD silicon carbide fibers have been shown to be anisotropic, with the radial strength being significantly lower than longitudinal tensile strength (Eldridge *et al.* 1993). Strength decreases in air when the fibers are heated to above 800 °C for long periods, most probably due to oxidation of the carbon-rich outer layers. At temperatures above 900 °C interfacial reactions between the tungsten core and mantle cause degradation of properties as explained above. At higher temperatures still grain growth in the silicon carbide mantle may be the cause of further falls in strength.

4. Conclusion

Chemical vapor deposition has produced filaments with extremely high mechanical properties. However their large diameters and hence lack of flexibility limits the forms into which they can be arranged. In addition, the necessity to fabricate the fibers individually, so that one reactor is necessary for each fiber made, means that production costs are high. These aspects of the fibers inevitably means that their uses are restricted to applications for which their remarkable properties are fully required. One such application could be the reinforcement of titanium alloys by silicon carbide CVD fibers in rotating parts of aircraft jet engines. The use of such reinforcements would allow the titanium to be used up to 600 °C whereas unreinforced the metal is limited by creep to a maximum of 450 °C.

Bibliography

Bhardwaj J, Krawitz A 1983 The structure of boron in boron fibres. *J. Mat. Sci.* **18**, 2639–49

Bunsell A R, Nguyen T T 1980 The radial strength of boron fibres and fiber splitting in boron–aluminium composites. *Compos. Fibre Sci. Tech.* **13**, 363–83

DiCarlo J A 1977 Time temperature stress dependence of boron deformation. *Composite Materials: Testing and Design.* ASTM, Philadelphia, PA, pp. 443–65

Eldridge J I, Wiening J P, Davison T S, Pindera M-J 1993 Transverse strength of SCS-6 silicon carbide fibers. *J. Am. Ceram. Soc.* **76** (12), 3151–4

Ericksen R H 1974 Room temperature creep and failure of borsic filaments. *Fibre Sci. Tech.* **7**, 173–83

Kreider K G, Prewo K M 1972 The transverse strength of boron fibers. *Composite Materials: Testing and Design.* ASTM STP 497. ASTM, Philadelphia, PA, pp. 539–50

Ning X J, Pirouz P 1991 The microstructure of SCS-6 SiC fiber. *J. Mater. Res.* **6** (10), 2234–48

Sattler M L, Kinney J H, Zywicz E 1992 The microstructures of SCS-6 and SCS-8 SiC reinforcing fibers. *Ceram. Eng. Sci. Proc.* **1992**, 227–37

Shatwell R A 1999 Fibre–matrix interfaces in titanium matrix composites made with sigma monofilament. *Mater. Sci. Eng. A.* **259**, 162–70

Vega-Boggio J, Vingsbo O 1976a Tensile strength and crack nucleation in boron fibres. *J. Mater. Sci.* **11**, 273–82

Vega-Boggio J, Vingsbo O 1976b Application of the Griffith criterion to fracture of boron fibres. *J. Mater. Sci.* **11**, 2242–6

Wawner F E 1988 Boron and silicon carbide fibres. In: Bunsell A R (ed.) *Fibre Reinforcement for Composite Materials.* Elsevier, Amsterdam, pp. 371–425

Whatley W J, Wawner F E 1985 Kinetics of the reaction between SiC (SCS-6) filaments and Ti-(6Zl-4V) matrix. *J. Mater. Sci. Lett.* **4**, 173–5

A. R. Bunsell
Centre des Matériaux, Evry, France

This Page Intentionally Left Blank

D

Designing with Composites

The key principles and issues in designing with composites are summarized in this article. More comprehensive treatments of this topic can be found elsewhere (e.g., US Department of Defense [USDD] (1999), Gillespie and Carlsson (1990), Lubin and Peters 1998). Given the dominance of fiber-reinforced polymer matrix composites in commercial applications, the discussion that follows focuses on this class of materials. Many of the same generic principles apply to other classes of composites (e.g., metal and ceramic matrices).

The heterogeneity and anisotropy of composites play an important role in the design process. A key concept is that the response of composite structures is determined by mechanisms and geometric features that exist at several different characteristic length scales and that contribute to the overall structural behavior (see *Composite Materials, Microstructural Design of**). These length scales include fiber diameter, ply thickness, laminate thickness, length of structural details (i.e., joints, thickness transitions, and cut-outs), and the overall structural dimensions. The ability to design at multiple length scales simultaneously via the choice of constituent materials, fiber architecture, and laminate stacking sequence offers the potential for highly optimized design solutions. However, this results in a highly-coupled, complex, and potentially expensive design process. Efficient and effective design with composites requires a good understanding of these issues, allowing for simplification as appropriate. In addition, the multiplicity of manufacturing routes available (see *Polymer Matrix Composites: Matrices and Processing**, *Metal Matrix Composites: Matrices and Processing**, *Ceramic Matrix Composites: Matrices and Processing**) and the strong dependence of the resulting mechanical properties on the manufacturing route employed, implies that design for manufacturability must also be considered.

1. *Overall Approach*

Composite structural design generally progresses in one or more steps from a preliminary design to a detailed design. The means by which this progression is achieved depends to a large degree on the complexity of the structure and the manufacturing routes to be employed. Overall approaches that have evolved range from analysis of the structural requirements and proceeding directly to the fabrication of prototype parts (e.g., tennis racquets or golf club shafts) to a more deliberate, multistep process (e.g., the "Building Block" approach used for aerospace vehicles). This methodology, illustrated in Fig. 1, combines fabrication, testing, and analysis at defined levels of design. Material "allowable stresses" and other parameters are determined through extensive testing and these parameters are transferred to progressively higher levels of structural complexity and/or scale (e.g., laminate/coupon, subelement, element, subcomponent, component, and full-scale structure) to allow design at these higher levels of complexity and/or scale. Analysis is used at each step, mainly to determine the state of stress and strain, while the "allowables" are passed on with appropriate knockdown factors and other empirical and semiempirical modifications to account for issues introduced at the next level of structural complexity. By including these empirically-based modifications at each level (or scale), the effects at the full-scale structure can be taken into account through comparison to past experience. It is important to note that for this approach to be successful it is required that the failure and deformation mechanisms remain the same as the design progresses from testing at one length scale to the next.

The efficiency of the design process has increased over time. In both prototype-driven design, and the building block approach, designers currently benefit from the pervasive use of computer-aided design and finite element analysis. The ability to explore rapidly a range of geometrical and material possibilities simultaneously is a key to maximizing the use of composites.

2. *Material System Selection*

The selection of a material system generally occurs during the preliminary design phase. It is a key decision since the choice of materials directly affects the overall structural configuration and manufacturing route to be used, which in turn have a strong influence on the cost of the system. Given the wide range of available materials, including composites, it is important that materials selection is carried out in a systematic manner (see *Fibrous Reinforcements for Composites: Overview**, *Glass Fibers**, *Fiber Strength**, *Metal Matrix Composites: Matrices and Processing**, *Ceramic Matrix Composites: Matrices and Processing**). The basic structural considerations for material selection are described in detail elsewhere (*Composite Materials, Microstructural Design of* *). It is worth noting that generally the selection of a composite material system is based on estimates of homogenized directional material properties, whereas the subsequent detailed design requires consideration of the material architecture (i.e., laminate, weave,

---

* Cross references marked by an asterisk are included in this volume.

# Cross references marked by a hash can be found by consulting the Encyclopedia of Materials: Science and Technology.

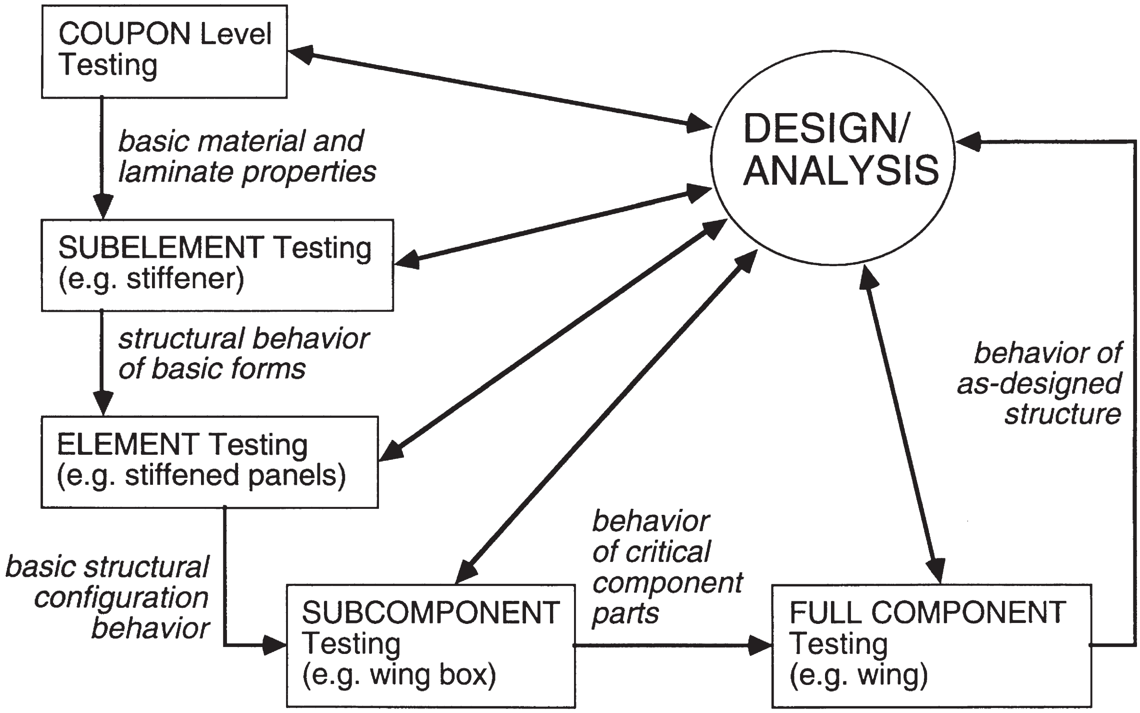

COUPON Level Testing
basic material and laminate properties
SUBELEMENT Testing (e.g. stiffener)
structural behavior of basic forms
ELEMENT Testing (e.g. stiffened panels)
basic structural configuration behavior
SUBCOMPONENT Testing (e.g. wing box)
behavior of critical component parts
FULL COMPONENT Testing (e.g. wing)
DESIGN/ ANALYSIS
behavior of as-designed structure

Figure 1
Flow chart for the "Building Block" approach for structural composite design/certification.

sandwich panel, etc.). Clearly it is important that the initial estimates are accurate for the achievable laminate (or other architecture) properties.

3. *Elastic Laminate Analysis*

Well-established analytical tools are available in the form of anisotropic elasticity and laminated plate theory (see *Laminates: Physical and Mechanical Behavior**). The homogenized ply properties are used as the initial material input to the analysis. A number of comprehensive references on these theories are available (e.g., Jones 1998, Reddy 1997). These analytical tools are available to the designer in the form of commercially available software packages and in embedded routines within finite element codes. The analytical tools previously noted allow for an accurate prediction of the elastic response of composite structures as well as the stress distribution on a ply-by-ply basis. Care must be taken in applying them to ensure that the modeling assumptions are appropriate for the structural configuration in question. For example, in some cases, either a full three-dimensional elasticity model must be developed whereas in others the three-dimensional effects can be incorporated into modified plate theory (Barbero *et al.* 1990).

4. *Design for Stiffness*

A number of special issues must be considered when designing with composites for stiffness-limited applications. In-plane stiffness anisotropy permits the structural stiffness to be locally tailored to meet the design requirements. In addition, composite laminates allow for the coupling between deformation modes, i.e., extension–shear coupling and coupling between bending and torsional deformations. This offers the potential for novel design solutions. However, these effects are generally avoided by using balanced, symmetric laminates in order to simplify the design process. This results in laminates with only orthotropic properties and no out-of-plane coupling. An additional important consideration is that the anisotropic thermal expansion at the ply level can result in thermal distortions in unsymmetric or imbalanced laminates. In practice choices are frequently further restricted to balanced, symmetric laminates with only 0°, ±45°, and 90° plies, and optimization is achieved by varying the fractions of plies of each orientation within this family. This restriction of the choice of laminate stacking sequence allows the use of graphical methods, such as carpet plots (Gillespie and Carlsson 1990), to identify appropriate stacking sequences. The choice of fiber direction roughly correlates with requirements for longitudinal, transverse, and shear loading.

5. *Design for Failure*

The ability to design against failure is less well developed. Numerous failure criteria have been proposed (e.g., Soden *et al.* 1998), but the resulting predictions are generally adjusted based on experimental data derived from mechanical testing together with "rules of thumb" accumulated from experience on previous projects and coupled with "factors of safety." The size of the test programs used in implementing the building block approach is largely determined by the uncertainties associated with predicting strength and probabilistic approaches have been developed (USDD 1999). Key issues in designing for failure include delamination, the role of free edges and stress gradients, the generally lower load carrying capacity in compression than tension, the effects of out-of-plane loading, and the importance of structural details. In addition damage tolerance and damage resistance requirements (particularly in terms of impact damage) have been regarded as a weakness of composites and require particular attention. These issues originate in the heterogeneous structure of composites and the resulting elastic anisotropy. Practical approaches to reducing the sensitivity of structures to these issues include: minimizing the number of adjacent plies with the same orientation (to suppress delamination); the use of internal ply drops to transition in thickness (to minimize free-edge stresses); and the use of through-thickness reinforcement (e.g., stitching) in areas susceptible to out-of-plane loading. The restriction of laminates to the $0/\pm45/90$ family also simplifies the process of designing against failure. No single analytical approach exists that can claim to allow for a sufficiently accurate predictive capability in order to permit simulation-based design against failure. This is a key limitation on the efficiency of the overall design process.

6. *Other Issues in Design*

In a similar situation to the case of designing against failure, several other issues in composite design are not amenable to analytical treatment. In particular, issues associated with the longevity and durability of composite parts have not proven to be amenable to analysis with sufficient accuracy to guide the design process. These include the effects of cyclic loading, moisture absorption, the effect of temperature, and the effects of active environments (see *Composite Materials: Environmental Effects**). As for the case of designing against failure, the approach taken has been to employ a combination of testing with heuristic knowledge gained from previous projects. A key difference for durability is that testing is not usually conducted at all levels in the building block approach. Typically, durability testing will be performed initially at either the coupon or subelement level and subsequently at the full-scale structural level. Damage tolerance (i.e., the ability of a damaged structure to continue to carry load) is generally only verified at the element level or higher. This is a matter of ensuring the cost-effectiveness of the process. The requirements for design for damage resistance, damage tolerance, and durability have led to the development of woven and stitched composites to limit the effects of impact loading, the development of more moisture-resistant matrix materials, and the development of matrices with higher temperature limits.

7. *Built-up Structure*

The design of large structures typically progresses according to the building block approach or analogous processes. As for the unit issues of designing against failure and for damage tolerance and durability, a combination of analysis, testing, and experience is brought to bear in the process. One effect of the reliance on the building block approach is the tendency for designs to gravitate towards solutions in which the large structure is built up from component parts using extensive secondary joining processes. This is apparently at odds with one of the promises originally made for composites, namely that relatively large structures can be made by net-shape forming methods, thereby reducing the number of parts and the number of unit processes required. In addition, the designers of large composite structures have generally not fully embraced the use of the efficient structural configurations and concepts originally envisioned for composites, such as the extensive use of sandwich panels, the use of cocured adhesive bonding, and the elimination of discrete stiffeners and mechanical fasteners. This conservative approach has arisen, in part, because of practical considerations and some adverse experiences. Honeycomb sandwich panels can be more sensitive to impact damage, and particularly moisture ingress, than discrete-stiffened structures. Mechanical fasteners are advantageous if a structure must be disassembled for inspection and repair. Cocuring large pieces of structure also has serious economic implications for the capital cost of the equipment required and the sensitivity to any manufacturing considerations.

8. *Summary*

Design approaches for composite structures are well-established and draw on 30 or more years of experience and analytical tool development. Generally these approaches have been successful in creating safe and effective designs. In both prototype-driven design, and the building block approach, designers currently benefit from the pervasive use of computer-aided design and finite element analysis. Further extension of this ability to include accurate predictive models and approaches for designing against failure and for durability and damage tolerance is a key requirement for increasing the efficiency of the overall design process.

Bibliography

Barbero E J, Reddy J N, Teply J 1990 An accurate determination of stresses in thick laminates using a generalized plate theory. *Int. J. Num. Methods Eng.* **29**, 1–14

Gillespie J W, Carlsson L A (eds.) 1990 *Delaware Composite Design Encyclopaedia.* Technomic, Lancaster, PA

Jones R M 1998 *Mechanics of Composite Materials.* Taylor & Francis, New York

Lubin G, Peters S T 1998 *Handbook of Composites*, 2nd edn. Chapman & Hall, London

Reddy J N 1997 *Mechanics of Laminated Composite Plates: Theory and Analysis.* CRC Press, Boca Raton, FL

Soden P D, Hinton M J, Kaddour A S 1998 A comparison of the predictive capabilities of current failure theories for composite laminates. *Comp. Sci. Tech.* **58**, 1225–54

US Department of Defense 1999 *Guidelines of Structural Materials.* Composite Materials Handbook, Vol 1. MIL-HDBK-17-E

S. M. Spearing and P. A. Lagace
Massachusetts Institute of Technology, Cambridge, Massachusetts, USA

E

Elastic Behavior of Cellular Solids

Cellular solids cover a wide range of multifunctional materials distinguished by low relative density. Familiar examples from nature and industry include cork, balsa wood, sponge, bread, and the flexible polyurethane foam that cushions our seats. More exotic foams manufactured from polymers, metals, and ceramics target high-technology aerospace and medical applications, and provide challenging opportunities to engineer new materials.

This article is concerned with the elastic properties of low-density cellular solids, which will also be referred to as foams. A micromechanical point of view provides a theoretical foundation for establishing structure–property–processing relationships, which guide the use and development of cellular solids. In general, stiffness and strength increase with the volume fraction, ϕ, of solid, $\phi = \rho/\rho_s$, which is synonymous with relative density; ρ is the foam density and the subscript s refers to the solid phase. The elastic constants exhibit strong dependence on microstructure when $\phi \ll 1$. Honeycombs and solid foams with open or closed cells, composed of slender struts or thin cell walls are used to illustrate the connection between macroscopic behavior, cell-level structure, and deformation mechanisms. All of the models used here are based on spatially periodic structures that involve a representative volume or "unit cell."

1. *Honeycombs*

The mechanics of "two-dimensional" honeycombs is a useful entry point into foam mechanics. The in-plane elastic constants of a regular hexagonal honey comb with $\phi \ll 1$ are

$$\frac{K}{E_s} = \frac{1}{4}\phi \quad \frac{E}{E_s} = \frac{3}{2}\phi^3 \tag{1}$$

where K is the bulk modulus and E is Young's modulus. These results illustrate a common theme for low-density cellular solids. A linear dependence of properties on ϕ is the signature of stretching of individual members, such as cell walls or struts. Higher order dependence indicates bending of individual members that causes a much softer response. The cubic dependence of E on ϕ corresponds to bending of thin cell walls. A honeycomb is relatively easy to compress in the plane because of the bending mechanism. Since $K \gg E$, the plane–strain Poisson's ratio $\nu \approx 1$. Equation (1) illustrates another important feature of cellular solids: the elastic response does not depend on cell size. In other words, self-similar microstructures have identical elastic behavior.

2. *Foams*

The microstructure of a solid foam forms in the liquid state, where gas bubbles grow into polyhedral cells that deform under the influence of flow imposed by the manufacturing process. This motivates using liquid foams as templates to create models of solid foams (see *Foams, Microrheology of**). The main feature of open-cell foams, which do not have cell walls, is a network of slender struts meeting at rigid joints. The Kelvin cell shown in Fig. 1(a) is a natural starting point because it defines the geometry of a perfectly ordered three-dimensional (3D) foam, which is analogous to the honeycomb in two dimensions (2D). The struts are modeled as beams with uniform cross-section composed of linear elastic material. There are three independent elastic constants because the structure has cubic symmetry. They can be calculated in closed form and show that the response is nearly isotropic (Warren and Kraynik 1997). The effective isotropic elastic constants evaluated for different strut shapes are

$$\frac{K}{E_s} = \frac{1}{9}\phi \quad \frac{E}{E_s} = C\phi^2 \tag{2}$$

The bulk modulus of this highly symmetric structure only involves the stretching mechanism. However, strut bending and torsion have comparable effects on E in this case and for more realistic foam microstructures. When $\nu_s = 1/2$, the parameter C ranges from 0.59–0.98 as strut shape varies from circular to Plateau border, the space between three mutually tangent circles. The Plateau border shape is found in real foams. Experimental data for different solid

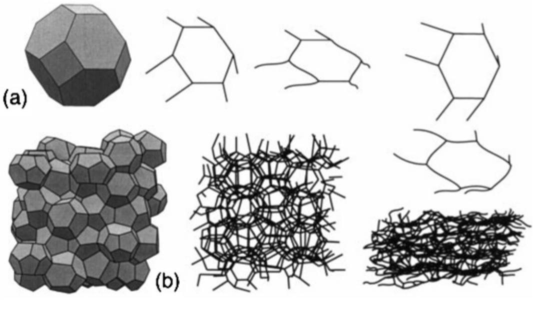


Figure 1
(a) Kelvin cell and (b) random monodisperse foam with 64 cells used as templates to produce the open-cell microstructures. "Pushing on a square and pushing on a hexagon" relate to the principal axes of the Kelvin cell. The deformed structures correspond to uniaxial compressive stress.

---

* Cross references marked by an asterisk are included in this volume.
# Cross references marked by a hash can be found by consulting the Encyclopedia of Materials: Science and Technology.

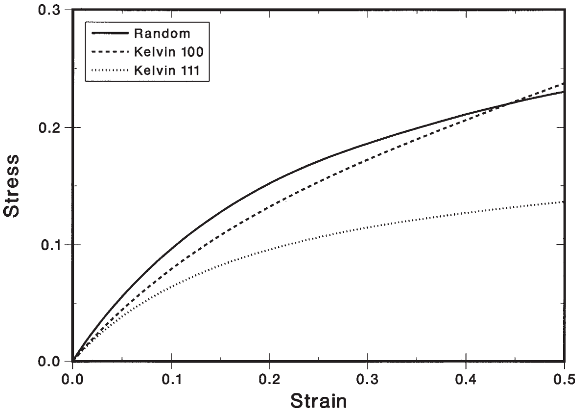


Figure 2
Stress–strain curves for uniaxial compression of the open-cell foams shown in Fig. 1 with struts of Plateau border shape and $\phi = 0.001$. The stress is scaled by $E_s\phi^2$. The two curves for the Kelvin foam correspond to pushing on a square face (100) or a hexagon (111); they have similar initial slopes but diverge at large compressive strains. ——, Random; ——, Kelvin 100; ⋯⋯, Kelvin 111.

foams can be found in Gibson and Ashby (1997); they use $C = 1$ to correlate data for open-cell foams.

Open-cell foams with disordered structure can also be patterned after liquid foams (see Fig. 1(b)). The resulting finite-element models are subjected to large uniaxial compression. The corresponding stress–strain curves for the Kelvin and random foam are shown in Fig. 2. The strut material is linear elastic so the nonlinearity stems entirely from large deformation of the microstructure. The Young's modulus for the random foam is about 30% larger than the Kelvin foam—disorder leads to a stiffer not softer response. Foam structures based on less realistic Voronoi cells are even stiffer.

Finite element analysis using shell elements to model a closed-cell Kelvin foam with $\nu_s = 0.49$ gives

$$\frac{K}{E_s} = 0.44\phi \quad \frac{E}{E_s} = 0.31\phi \tag{3}$$

Both properties scale with ϕ implicating the stretching mechanism and giving the stiffest response of any microstructure considered here. The foam has a Poisson's ratio $\nu = 0.38$. These results are typical of those obtained for more complicated foam structures.

Equation (3) neglects the effect of gas pressure P_g in the closed cells. Accounting for the gas would add the term P_g/E_s to the right-hand side of the bulk modulus. This contribution can be important when the solid phase is very soft or under large compression.

3. *Concluding Remarks*

This short survey can only hint at the rich mechanical behavior of real foams. Equations (2) and (3) are useful guidelines, but real foams often have intermediate microstructure and properties caused by the presence of cell walls and thick cell edges where struts form. When ϕ is a few percent, the struts and joints in open-cell foams can be quite thick, so asymptotic results developed for $\phi \ll 1$ must be kept in perspective. Other complications and imperfections abound.

Many applications of foams involve crushing. The stress–strain curves in Fig. 2 suggest but do not exhibit the extended plateau and lack the eventual sharp increases in stress that are often seen in laboratory measurements. The plateau defines material strength and is caused by "buckling" of the microstructure and strain localization, which have been analyzed for honeycombs (Papka and Kyriakides 1998) but not foams. This is just one example of the challenge and opportunity in foam mechanics.

Bibliography

Christensen R M 2000 Mechanics of cellular and other low-density materials. *Int. J. Solids Struct.* **37**, 93–104

Gibson L J, Ashby M F 1997 In: *Cellular Solids: Structure and Properties*, 2nd edn. Cambridge University Press, Cambridge

Kraynik A M, Neilsen M K, Reinelt D A, Warren W E 1999 Foam micromechanics: structure and rheology of foams, emulsions, and cellular solids. In: Sadoc J F, Rivier N (eds.) *Foams and Emulsions, Proc. School on Foams, Emulsions, and Cellular Materials*, Cargese, France, May 12–24, 1997. Kluwer, Dordrecht, The Netherlands, pp. 259–86

Kraynik A M, Warren W E 1994 The elastic behavior of low-density cellular plastics. In: Hilyard N C, Cunningham A C (eds.) *Low-Density Cellular Plastics*. Chapman and Hall, London, pp. 187–225

Papka S D, Kyriakides S 1998 In-plane crushing of a polycarbonate honeycomb. *Int. J. Solids Struct.* **35**, 239–67

Warren W E, Kraynik A M 1997 Linear elastic behavior of a low-density Kelvin foam with open cells. *J. Appl. Mech.* **64**, 787–94

A. M. Kraynik and M. K. Neilsen
Sandia National Laboratories, Albuquerque,
New Mexico, USA

Elastic Behavior of Composites

Other articles in this encyclopedia are devoted to composites covering reinforcing fibers, matrices of different natures (polymeric, metal, and ceramic), their processing, mechanical behavior at the micro- and macromechanical scale, designing principles, and miscellaneous applications of composite structures. Inevitably there is some degree of overlap, but the primary aim of this article is to cover elastic properties and to emphasize the ways in which composites differ from other materials. It is for this reason that the following issues are concentrated on: the lamina as the building block with four elastic constants characterizing the stress/strain response, the directional nature of composite stiffness values, the coupling between normal stresses and shear strains and vice versa, the effect of stacking on the laminate stiffness, the unusual Poisson ratios outside the limit usual for isotropic materials, and the coupling phenomena between extension and bending for unsymmetrical laminates.

1. *Lamina Stiffness Characteristics*

The building block of a composite is called the lamina consisting of parallel, continuous, stiff, and strong fibers embedded in a matrix. The lamina properties are dictated by the constituent properties, typically by Young's modulus, E, and Poisson's ratio, v, and their concentrations, e.g., by the fiber volume fraction, v_f. Micromechanical modeling (see *Composites, Physical Properties of**) enables the lamina properties to be derived.

1.1 The Basic Stiffness Characteristics

As a lamina is transversely isotropic, there are two principal directions associated with extreme properties. These are the longitudinal direction parallel with fibers and the transverse direction perpendicular to the fibers. The former is denoted by the subscript 1, the latter by the subscript 2. On loading a lamina in either principal direction, its behavior is orthotropic. In other words, extension or compression is not accompanied by any distortion of angles.

Assuming further that the lamina finds itself under the plane stress situation, i.e., its thickness is negligible in comparison with the width and length, the stress/strain behavior involves two normal stresses, σ_1, σ_2, and one shear stress, σ_{12}, related to the corresponding normal (ε_1 and ε_2) and the shear strains, ε_{12}, by

$$\begin{aligned} |\sigma| &= |Q| * |\varepsilon| \\ |\varepsilon| &= |S| * |\sigma| \end{aligned} \tag{1}$$

where the stiffness matrix, $|Q|$, and the compliance matrix, $|S|$, are of the type

$$|Q| = \begin{vmatrix} Q_{11} & Q_{12} & 0 \\ Q_{12} & Q_{22} & 0 \\ 0 & 0 & Q_{66} \end{vmatrix}; \quad |S| = \begin{vmatrix} S_{11} & S_{12} & 0 \\ S_{12} & S_{22} & 0 \\ 0 & 0 & S_{66} \end{vmatrix} \tag{2}$$

The four components of these matrices are related to the so-called engineering constants of the lamina, namely to the longitudinal, transverse, and shear moduli, E_1, E_2, and G_{12}, supplemented by Poisson's ratio, v_{12}. Apart from the fact that these can be derived micromechanically, they can be also measured using a spectrum of dedicated tests (Tarnopolskii and Kincis 1985). The relationships for the elements of the compliance matrix are simpler and are:

$$\begin{aligned} S_{11} &= \frac{1}{E_1} \quad S_{22} = \frac{1}{E_2} \\ S_{66} &= \frac{1}{G_{12}} \quad S_{12} = -\frac{v_{12}}{E_1} \end{aligned} \tag{3}$$

The lamina can be considered to be the composite entity of the highest anisotropy; extremely strong and stiff in the fiber direction, weak and soft in the perpendicular (matrix) directions. Plytron lamina based on polypropylene matrix reinforced with 50% (by weight) of E-glass fibers will be used in this article for illustrations of all points made. This is for two reasons: (a) this material type has already found its role in automotive applications; and (b) in the author's opinion Plytron lamina is likely to become one of the major composites used in the future. Engineering constants of the Plytron lamina have been determined as $E_1 = 21.7$ GPa, $E_2 = 3.34$ GPa, $G_{12} = 1.32$ GPa, and $v_{12} = 0.35$.

1.2 Stiffness Characteristics of Lamina Loaded in an Off-axis Direction

If the lamina is subjected to a load at an arbitrary angle $0 < \Theta < 90°$, the stress/strain situation is much more complex because the axes related to the material symmetry system does not coincide with the coordinate system dictated by loading. Thus both the stress and strain components denoted 1, 2, and 12 and derived as above have to be transformed into x, y, and xy equivalents with the x-direction determining that of a load. The stress/strain behavior in the new coordinate system relevant to the structural response then involves the so-called reduced transformation stiffness/compliance matrices $|\underline{Q}|$ and $|\underline{S}|$ of the following format:

$$\begin{aligned} \begin{vmatrix} \sigma_x \\ \sigma_y \\ \sigma_{xy} \end{vmatrix} &= |\underline{Q}| \begin{vmatrix} \varepsilon_x \\ \varepsilon_y \\ \varepsilon_{xy} \end{vmatrix} = \begin{vmatrix} \underline{Q}_{11} & \underline{Q}_{12} & Q_{16} \\ \underline{Q}_{12} & \underline{Q}_{22} & Q_{26} \\ \underline{Q}_{16} & \underline{Q}_{26} & Q_{66} \end{vmatrix} \begin{vmatrix} \varepsilon_x \\ \varepsilon_y \\ \varepsilon_{xy} \end{vmatrix} \\ \begin{vmatrix} \varepsilon_x \\ \varepsilon_y \\ \varepsilon_{xy} \end{vmatrix} &= |\underline{S}| \begin{vmatrix} \sigma_x \\ \sigma_y \\ \sigma_{xy} \end{vmatrix} = \begin{vmatrix} \underline{S}_{11} & \underline{S}_{12} & \underline{S}_{16} \\ \underline{S}_{12} & \underline{S}_{22} & \underline{S}_{26} \\ \underline{S}_{16} & \underline{S}_{26} & \underline{S}_{66} \end{vmatrix} \begin{vmatrix} \sigma_x \\ \sigma_y \\ \sigma_{xy} \end{vmatrix} \end{aligned} \tag{4}$$

The elements $\underline{Q}_{ij}$ and $\underline{S}_{ij}$ can be derived from the elements of the stiffness/compliance matrices Q_{ij} and S_{ij} and trigonometric function of the loading angle Θ

between the fiber and loading directions:

$$\underline{Q}_{ij} = q(ij)[Q_{11}, Q_{22}, Q_{12}, Q_{66}, \sin^n\Theta, CP\cos^m\Theta]$$

$$\underline{S}_{ij} = s(ij)[S_{11}, S_{22}, S_{12}, S_{66}, \sin^n\Theta, CP\cos^m\Theta] \quad (5)$$

where $q(ij)$ and $s(ij)$ are explicit functions for the element ij. In addition, the elements $\underline{S}_{ij}$ and $\underline{Q}_{ij}$ can be related to the stiffness characteristics relevant to the structural response of the lamina, i.e., E_x (loading direction), E_y (direction perpendicular to the load), G_{xy}, and ν_{xy} (shear modulus and the Poisson ratio in the lamina plane).

Figure 1 shows angular variation of these stiffness characteristics normalized by the appropriate engineering constants of the Plytron lamina. E_x/E_1 is seen to be a decreasing monotonic function of the loading angle Θ. The dependence of G_{xy} is a symmetrical function of Θ, $G_{xy}(\Theta) = G_{xy}(90-\Theta)$, passing through a maximum at $\Theta = 45°$. The angular dependence of ν_{xy} satisfies the boundary conditions $\nu_{xy}(0)/E_1 = \nu_{xy}(90)/E_2$ and exhibits a maximum in the region of low loading angles. The area under the $E_x(\Theta)$ dependence represents composite potential contained in the lamina: it is relatively low for thermoplastic matrices reinforced with glass (the case of Plytron) and high for laminas containing a large concentration of high-modulus carbon fibers.

In conclusion, while the stress/strain behavior of the lamina loaded in the principal direction (orthotropic behavior) is described by four elements Q_{ij} (or S_{ij}), the anisotropic response of the lamina loaded at any other angle requires six elements $\underline{Q}_{ij}$ (or $\underline{S}_{ij}$). At the same time, they can be derived from the four basic engineering constants characterizing the lamina.

Two additional elements, $\underline{Q}_{16}$ and $\underline{Q}_{26}$ ($\underline{S}_{16}$ and $\underline{S}_{26}$, respectively) are associated with another feature of the elastic response of composites, namely with the coupling between normal strains and shear stresses and vice versa. As this behavior is unique to composites, the coupling effects are further expanded in the following section.

1.3 Coupling Between Normal Stresses and Shear Strains

Tensile deformation of isotropic materials is associated with extension in the direction of the load and contraction in the lateral direction. If a lamina is loaded in an off-axis direction, deformation is anisotropic as it also involves changes in angles. This is owing to coupling of normal stresses with shear strains and shear stresses with normal strains. The extents of these couplings are quantified by the so-called coupling coefficients, ν_{ij}, defined as

$$\nu_{ij} = \frac{\underline{S}_{ij}}{\underline{S}_{ii}} \quad (6)$$

Figure 2 shows the angular dependencies of ν_{16} and ν_{26} for the Plytron lamina and compares them with the dependence of Poisson's ratio, ν_{xy}. While Poisson's ratio attains positive values only, the variation in the coupling coefficients is more dramatic. The functions involve both slightly positive and large negative values. In absolute terms, the effects appear more significant than those characterized by Poisson's ratio. There is an axial symmetry of 16 and 26 couplings: $\nu_{16}(\theta) = \nu_{26}(90-\theta)$.

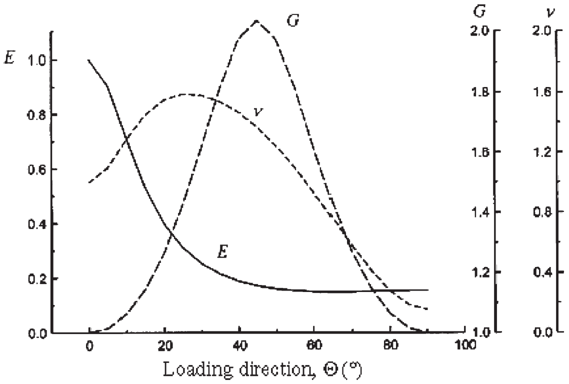


Figure 1
Angular dependence of the off-axis stiffnesses normalized on the lamina characteristics of Plytron. Normalization: $E = E(\theta)/E_1$, $G = G(\theta)/G_{12}$, $\nu = \nu(\theta)/\nu_{12}$.

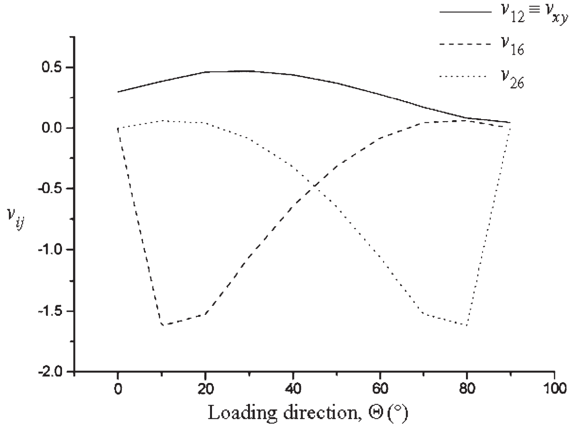


Figure 2
Angular dependence of the coupling coefficients and the Poisson ratio for the Plytron lamina.

It is this coupling owing to a combination of lamina anisotropy with geometric factors that leads to the state of combined stresses and that makes composites unique in their performance. The analysis of this unusual material behavior is simple although the mathematical apparatus needed is more involved than that which suffices for isotropic materials. Attention to all details is required so that the coupling effects may be exploited fully as an opportunity rather than ignored and their existence considered as a composite liability. The discussion of laminates reinforces this theme.

2. *Laminates*

2.1 Effective Stiffnesses in the Lamina Plane and the Coupling Coefficients

The analysis outlined in Sect. 1.2 can be easily extended to the behavior of laminates, i.e., structures containing a certain number, n, of laminas (plies) of thickness t stacked together in the direction z perpendicular to the laminate plane to form a monolithic structure of overall thickness $h = tn$. Furthermore, for any lamina i the fiber direction is offset by the angle Θ_i with respect to the loading direction and the space occupied by that is between z_i and z_{i-1}. The upper plane of the top lamina is thus situated at z_0 and the lower plane of the bottom lamina at z_n.

As a laminate is a layered structure with each ply characterized by its own stiffness, $E(\Theta_i)$, the whole construction is monolithic (no mutual slippage of adjacent plies allowed), the "state equation" of laminates involves another coupling, namely that between extension and bending (Reissner and Stavsky 1961):

$$\begin{aligned} |N| &= |A||\varepsilon^0| + |B||\kappa| \\ |M| &= |B||\varepsilon^0| + |D||\kappa| \end{aligned} \tag{7}$$

A thermometer based on a bimetallic element serves as a good example of the principles involved.

With reference to *Laminates: Physical and Mechanical Behavior**, $|A|$ is defined as the extensional matrix, $|B|$ as the coupling matrix, and $|D|$ the bending matrix. The force matrix $|N|$, the moment matrix $|M|$, the middle-stack strain matrix $|\varepsilon^0|$, and the curvature matrix $|\kappa|$ are all obtained by appropriate integration of matrices defined by Eqn. (4) over the laminate thickness h. A discussion of the elastic behavior of laminates deserves a more detailed analysis of the matrices $|A|$, $|B|$, and $|D|$. These are 3×3 matrices based on six elements of the same indices as those contained in the reduced matrix $|\underline{Q}|$ (Eqn. (4)). Any element ij can be derived by summing up contributions of all plies from the corresponding $\underline{Q}_{ij}$ and taking into account the actual position of the kth ply in the stack:

$$\begin{aligned} A_{ij} &= \sum_{k=1}^{n} (\underline{Q}_{ij})_k (z_k - z_{k-1}) = t \sum_{p=1-n/2}^{n/2} (\underline{Q}_{ij})_p \\ B_{ij} &= \frac{1}{2} \sum_{k=1}^{n} (\underline{Q}_{ij})_k (z_k^2 - z_{k-1}^2) = \frac{t^2}{2} \sum_{p=1-n/2}^{n/2} (\underline{Q}_{ij})_p [2p-1] \\ D_{ij} &= \frac{1}{3} \sum_{k=1}^{n} (\underline{Q}_{ij})_k (z_k^3 - z_{k-1}^3) \\ &= \frac{t^3}{3} \sum_{p=1-n/2}^{n/2} (\underline{Q}_{ij})_p [3p^2 - 3p + 1] \end{aligned} \tag{8}$$

where p is the ply index expressing position of a given ply in another, more convenient, manner. The index $p = 0$ identifies the first ply below the midplane, $p = n/2$ the top ply, and $p = 1-n/2$ the bottom ply. The weighting factor in the square brackets implies that if the stack is symmetrical with respect to the midplane the coupling matrix $|B|$ is zero. This means that extensions are decoupled from bending, i.e., an exerted force causes exclusively extensions and moments are responsible for bending.

Finally, defining the matrices $|A^*|$ and $|D^*|$ to be the inverse to matrices $|A|$ and $|D|$, all the laminate stiffness characteristics can be expressed as

$$\begin{aligned} E_x &= \frac{1}{A*_{11}} h \\ \nu_{12} &= -\frac{A*_{12}}{A*_{11}} \end{aligned}$$

The effective moduli are seen to be arithmetic averages of stiffnesses of the constituent plies. However, two-dimensional composite properties expressed by the coupling coefficients differ significantly from two-dimensional properties of conventional isotropic materials. Coupling coefficients can be shown to be large and their origins not always obvious. Furthermore, they can be stated to represent the real potential of composites.

2.2 Some Important Symmetrical Stacks

(*a*) *Cross-plied laminates.* These are obtained by stacking 0° and 90° plies based on an identical lamina with a ply, p, being counterbalanced by the ply of the same orientation. Loading in either fiber direction results in an orthotropic laminate response and the coupling coefficients ν_{16} and ν_{26} are zero. An effective way of controlling the laminate stiffness is to vary the so-called ply ratio, M, i.e., the fraction of 0° plies in the stack. This is illustrated in Fig. 3 showing the dependence of the normalized (on the appropriate

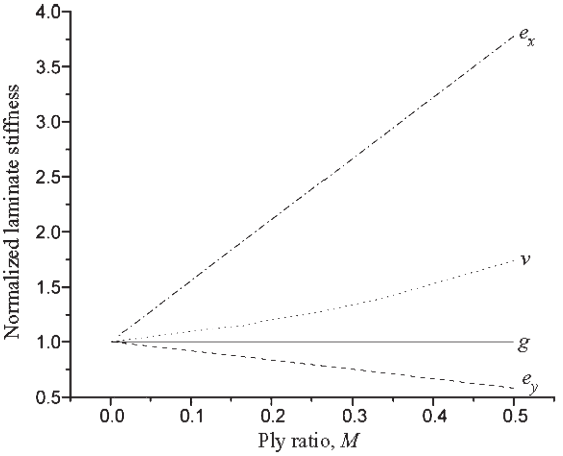


Figure 3
Dependence of the stiffness characteristics on the construction of cross-plied laminates based on the Plytron lamina. Normalisation: $e_x = E_x/E_1$, $e_y = E_y/E_2$, $g = G_{xy}/G_{12}$, $v = v_{xy}/v_{12}$.

engineering constants of the lamina) stiffnesses in the longitudinal and transverse directions, shear modulus, and Poisson's ratio in the laminate plane. As the ply ratio increases, the $e_x = E_x/E_1$ dependence linearly increases, $e_y = E_y/E_2$ linearly decreases, and $g = G_{xy}/G_{12}$ remains independent of M. There is a marginal nonlinear increase in the relative Poisson's ratio $v = v_{xy}/v_{12}$ where 90° plies play a constraining effect. The simplest construction of the cross-plied laminate is the {0/90/90/0} and associated with $M = 0.5$.

Loading a cross-plied laminate in a direction not coinciding with that of the fibers in plies is associated with an anisotropic response with $v_{16} \neq 0$ and $v_{26} \neq 0$. Figure 4 shows the variation of v_{16} with the loading angle $90° \geqslant \Theta \geqslant 0°$ and the ply ratio $1 \geqslant M \geqslant 0$. The coupling coefficient attains both the positive and negative values far exceeding unity. In addition, the variations are rather abrupt. This is particularly true for loadings at large angles and for low M values, i.e., a drastic change in the behavior owing to addition of a single 0° ply to the originally unidirectional laminate consisting of 90° plies.

(*b*) *Angle-plied laminates.* The symmetric stack consists of $+\omega$ and $-\omega$ plies. The absolute value of the ply angle, ω, controls the laminate stiffness values as illustrated in Fig. 5 showing the angular dependence of the four stiffness characteristics of angle-plied laminates $E_x(\omega)$, $E_y(\omega)$, $G_{xy}(\omega)$, and $v_{xy}(\omega)$ normalized on the stiffness characteristics representing the behavior of the uniaxial construction loaded in the corresponding $\Theta \equiv \omega$ direction. For Plytron laminates, the alternate stacking of plies enhances the axial and transverse stiffness E_x and E_y by more than 80% and leads to a broader G_{xy} peak being 120% above the maximum level exhibited by the uniaxial construction. Poisson's ratio is seen to involve positive values exceeding the upper bound for isotropic materials of 0.5. In addition, $\omega = 70°$ construction is associated with negative values. With the loading directions bisecting the ply angles, the laminate behavior is orthotropic with $A_{16} = A_{26} = 0$.

A much wider range of Poisson's ratios is obtained when an angle-plied laminate is loaded along a direction different from the bisector of the two ply angles. This is illustrated in Fig. 6 showing a contour map of Poisson's ratio as a function of the ply angle, ω, and the load direction, Θ. Variation in Poisson's ratio conforms to the diagonal symmetry with the negative values attained by the 45° laminates at $\Theta = 30°$ and 60°. The maximum positive values of Poisson's ratio are observed for 30° laminates at $\Theta = 0°$ and 60° laminates at $\Theta = 90°$, respectively.

(*c*) *Quasi-isotropic laminates.* A very popular laminate construction is achieved when the ply angle of two subsequent plies is rotated by π/n. The simplest laminate of this series of fourteen possible is $\{0/60/-60\}_s$ and the most popular is $\{0/45/90/-45\}_s$. The features of this laminate type are that E_x and E_y are independent of the loading direction and that coupling is absent by virtue of $A_{16} = A_{26} = 0$. The laminate stiffness cannot be altered unless a different building lamina is used. Thus this type of laminate represents the antipode to the unidirectional construction. It possesses minimum anisotropy, offers lowest performance, and may be qualified for lightweight replacement goals rather then for engineering ambitions. In composite technology, quasi-isotropic laminates represent the upper bound on weight and may be considered to be a useful starting point for a possible composite application with a possibility to optimize and to improve the design by enhancing anisotropy.

2.3 *Asymmetric (unbalanced) Laminates*

If the midplane symmetry is allowed to be absent, the elastic behavior of general laminates described by Eqn. (7) involves coupling between extension and bending and thus the full analysis that quantifies the interaction among forces, moments, extensions, and curvatures is required. The procedure is more cumbersome; nevertheless, the same methodology is used as already described.

An important issue that concerns this (general) laminate type is the way of controlling the coupling matrix $|B|$. Concentrating on laminates containing an even number of plies stacked alternately, the structural parameter of the laminate construction that controls coupling between extension and bending is

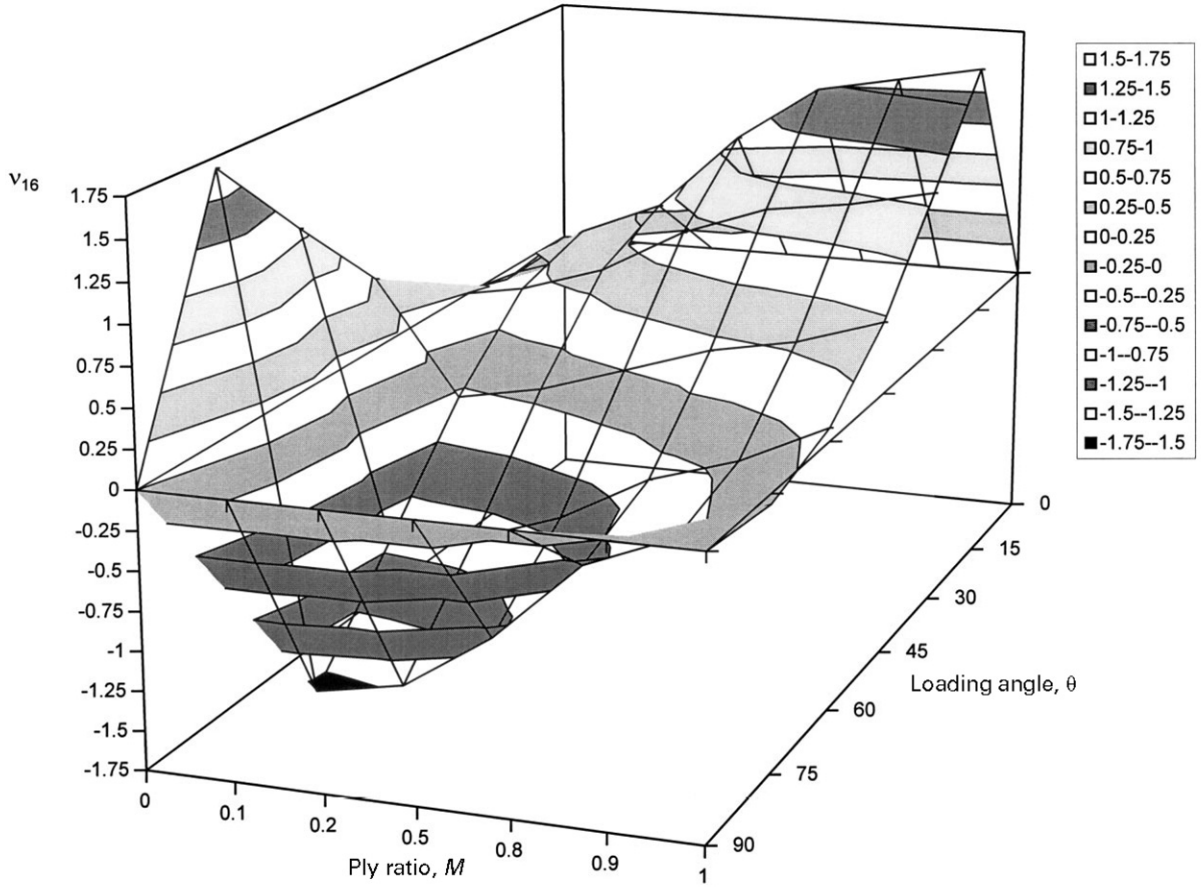


Figure 4
Variation in the coupling coefficient with the loading direction and the ply ratio of the cross-plied laminates based on the Plytron lamina.

the number of plies, n, as illustrated in Figs. 7 and 8. For the angle-plied laminates (Fig. 7), the extent of coupling quantified by B_{16}/A_{11} is correlated with the ply angle and the number of plies, n. At a fixed value of n characterizing this laminate type (so-called antisymmetric), the maximum coupling is achieved for the $+45/-45$ laminates. Coupling is seen to be reduced by the number of plies and thus the construction $\{+45/-45\}_T$ represents the case of the strongest coupling. For the cross-plied laminates (Fig. 8), the extent of coupling proportional to B_{11}/A_{11} is correlated with the structural parameter n and the lamina anisotropy expressed as the ratio E_2/E_1. For a given lamina, e.g., for that of Plytron characterized by $E_2/E_1 = 0.154$, the maximum coupling is again achieved for the construction $\{0/90\}_T$. If the laminate construction is fixed, the extent of coupling increases with increasing lamina anisotropy.

It is this combination of the material choice (lamina anisotropy) and the laminate construction that gives unique possibilities to a composite designer of structures such as the fan blade for which the centrifugal stress and the aerodynamic twist occur at the same time. Composites thus start fulfilling the role of tailored functional materials and this attribute defines their unique place in future technologies.

3. *Concluding Remarks*

Features of composites that make this material type different from all the others have been described. The transversely isotropic lamina acts as the building block and four engineering constants are required to describe its stress/strain behavior: longitudinal and transverse stiffnesses, E_1 and E_2, in the principal directions, and shear modulus, G_{12}, and Poisson's ratio, ν_{12}, in the lamina plane all controlled by the constituent properties and their mutual concentrations. Lamina stiffness characteristics have been shown to

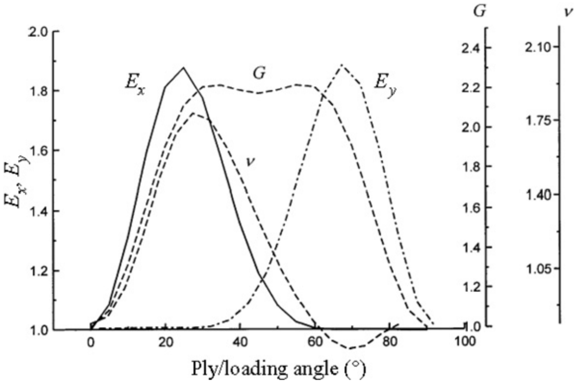


Figure 5
Angular dependence of the normalized stiffnesses for the Plytron laminates. Normalization: $E_x = E(\omega)/E(\theta)$, $E_y = E(\omega)/E(\theta)$, $G = G(\omega)/G(\theta)$, $\nu = \nu(\omega)/\nu(\theta)$.

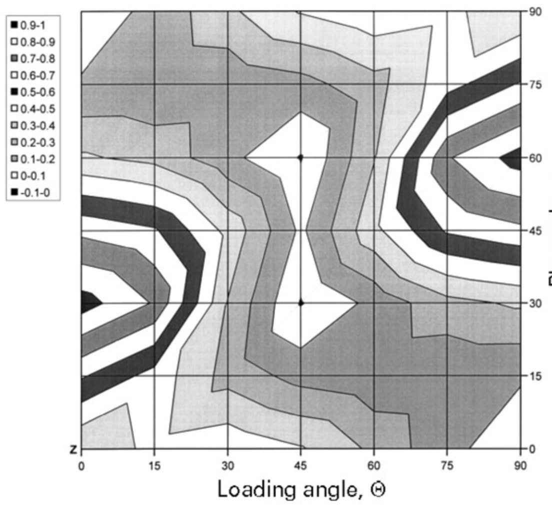


Figure 6
Variation in the Poisson ratio, ν_{xy}, with the loading direction for different angle-plied laminates based on the Plytron lamina.

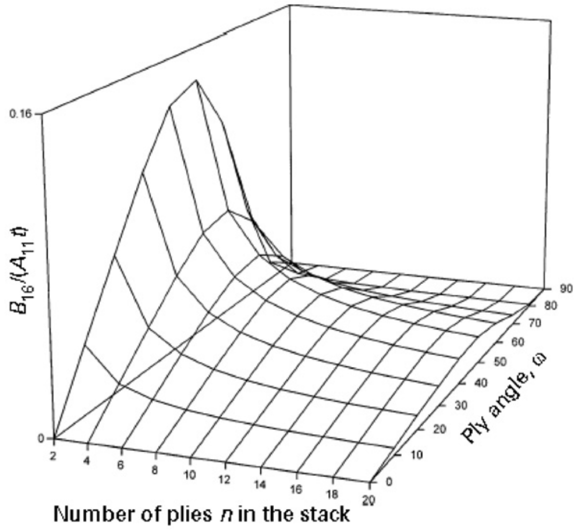


Figure 7
Construction of unbalanced angle-plied laminates and coupling between extension and bending.

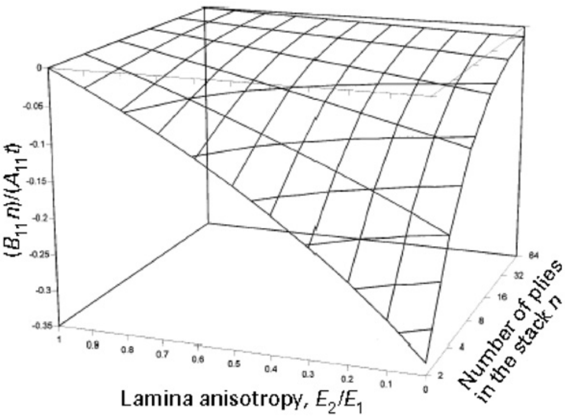


Figure 8
Effect of the construction of unbalanced cross-plied laminates on the coupling between extension and bending.

be dependent on direction and, in addition, there is a coupling between normal stresses and shear strains, and vice versa unless the loading happens in one of the principal directions. Stacking several laminas to form a monolithic laminate has been demonstrated to be the way towards simple structural elements exhibiting both Poisson's ratios and shear couplings not attainable by means of conventional materials. If the stacking arranges for midplane symmetry, the laminate behavior is relatively simple as the laminate extension is decoupled from the laminate bending. With the laminate midplane absent, the coupling matrix $|B|$ is nonzero and the choice of a proper lamina, combined with its orientation and the way of stacking, opens ways towards structural responses in which composites behave as functional materials, which will be used in future technologies.

So in the terms of the concepts developed in this article the elastic behavior of composites is to be understood such that the lamina formulation controls the stiffness and compliance matrices $|Q|$ and $|S|$, the lamina orientation their reduced equivalents $|\underline{Q}|$ and $|\underline{S}|$, and the stacking sequence the three relevant matrices $|A|$, $|B|$, and $|D|$. This universal procedure then allows either a simple stress/strain response to be orchestrated by eliminating anisotropy and working with quasi-isotropic laminates with the midplane symmetry, or, alternatively, to exploit fully the possible interaction of forces, moments, extensions, and curvatures to enter the domain of new materials.

See also: Laminates: Physical and Mechanical Behavior*

Bibliography

Agarwal B D, Broutman L J 1980 *Analysis and Performance of Fiber Composites*. Wiley, New York

Jones R M 1975 *Mechanics of Composite Materials*. Scripta Book Company, Washington, DC

Pipes R B, Pagano N J 1970 Interlaminar stresses in composite laminates. *J. Composite Mater*. **4**, 538–48

Reissner E, Stavsky Y 1961 Bending and stretching of certain types of heterogeneous aelotropic plates. *J. Appl. Mech*. **9**, 402–8

Springer G S 1981 *Environmental Effects on Composite Materials*. Technomic, Westport, CT

Tarnopolskii Yu M, Kincis T 1985 *Static Test Methods for Composites*. Nostrand Reinhold, Melbourne, Australia

Tsai S W, Hahn H T 1980 *Introduction to Composite Materials*. Technomic, Westport, CT

Vinson J R, Sierakowski R L 1986 *The Behaviour of Structures Composed of Composite Materials*. Kluwer, Dordrecht, The Netherlands

A. Cervenka
University of Manchester Institute of Science and Technology, UK

Elastic Structures in Design

Structures may be broadly defined as engineering artifacts that support loads. The loading causes the material within the structure to become stressed, enabling loads to be transferred from where they are applied (including the weight of the structure itself) to the surroundings. The term "structures" usually brings to mind bridges and buildings. Taking a broader perspective than just civil engineering however, it can be seen that structures come in all shapes and sizes. For example, most modes of transport involve structural design, e.g., aircraft wings, car bodies, and bicycle frames. This also applies to many other familiar load-bearing products, such as chairs, tables, and sports products (vaulting poles, tennis racquets, oars, etc.).

One design characteristic of any structure is its stiffness—that is, the deflection per unit of loading. Stiffness is controlled by both the geometry of the structure (the size and shape of its components) and the material used. Most structures are designed to remain elastic in service, so the maximum stresses must be analyzed to avoid failure by fracture or plastic collapse. Structural design can sometimes be concerned with large plastic deformations, e.g., for energy absorption in automotive crumple zones and crash barriers.

The subject of this article is elastic design and its interaction with material selection. The emphasis is on design for stiffness, although elastic design can have other aspects. For example, springs are designed to store elastic strain energy, while other structures are limited by vibration, either to reduce noise and interference (as in a vehicle engine or suspension), or because the structure is designed to generate sound (as in a musical instrument).

1. Elements of Elastic Design

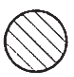
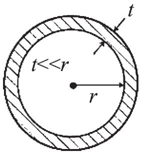
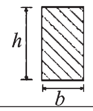
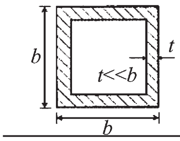
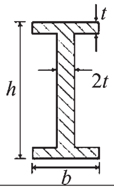
The key material property in stiffness-limited design is Young's modulus, E (or the related shear modulus, G, for loading involving shear). However, an important aspect of determining the deflection per unit load is also the mode of loading: tension, bending, torsion, or compression. Each mode is characterized by particular structural elements (e.g., beams in bending, shafts in torsion), and the components of elastic structures can always be approximated to some combination of these. For example, the main down-tube of a diamond bicycle frame primarily experiences torsion from the pedaling action of the cyclist (along with some tension from the overall triangulated shape of the frame, and bending due to it being rigidly jointed).

For structures in which forces are carried axially, such as cables or the members in a truss, optimizing stiffness is a straightforward design problem. The elastic deflection of the overall structure is simply related to the extensions in its members, and for axial loading these extensions scale inversely with the cross-sectional area of the component and Young's modulus. In this case, the cross-sectional shape has no influence.

Bending is more complex and usually dominates deflection-limited design (since much smaller loads are required to produce distortion in bending than by axial loading). Closely related is *buckling*—an instability caused by axial compression on a structural member, when the member deflects sideways and is essentially loaded in bending. Buckling is discussed in *Elastic Instability*# and is not considered further here.

The loading on a beam in bending generates *bending moments* along its length. In classical theory, the moments cause a linear stress distribution from tension to compression across the section (with the "neutral axis" being the axis where the stresses change from tension to compression). The stresses

Table 1
Second moments of area and shape factors for standard structural sections (after Ashby 1999).

Section shape (*equal areas*)	**Second moment of area**	**Shape factor in elastic bending, Φ**	**Typical values of shape factor, Φ**
	$\pi r^4/4$	1	1
	$\pi r^3 t$	r/t	20–40
	$bh^3/12$	$(\pi/3)\times(h/b)$	2–10
	$\frac{2}{3}b^3t$	$0.52\times(b/t)$	10–20
	$\frac{1}{6}h^3t\left(1+\frac{3b}{h}\right)$	$0.52\times\frac{h}{t}\times\frac{(1+3b/h)}{(1+b/h)^2}$	30–80

cause curvature of the beam, with resulting deflection. The overall geometry also depends on the way the beam is supported, e.g., cantilever or end-supported, rigidly clamped or free to rotate, etc.

In the analysis of stresses and deflections in a beam, the most important characteristic is its cross-sectional shape. For a given applied bending moment, the section shape governs the slope of the stress gradient across the section and thus the curvature in a given material. Shape is characterized by the second moment of area, I, defined as

$$I = \int y^2 \mathrm{d}A$$

where y is the distance of the elemental cross-sectional area dA from the neutral axis. Table 1 summarizes formulas for I in standard structural shapes. Efficient shapes in bending are those that maximize I for a given cross-sectional area of material, such as I-beams. In these shapes most of the cross-section is located as far from the neutral axis as possible (large y).

A useful parameter has been defined to quantify the effect of shaping a given cross-sectional area for bending. This dimensionless "shape factor," Φ, is defined by

$$\Phi = \frac{4\pi I}{A^2}$$

The constant 4π is introduced such that $\Phi = 1$ for a solid cylinder, i.e., Φ measures the ratio of I for a given shape to I for the same area made into a solid cylinder. Typical values for Φ are included in Table 1.

For simple idealized loadings, the maximum deflection, δ, is given by

$$\delta = C_1 \frac{WL^3}{EI}$$

where the constant C_1 only depends on the detail of the way the load is distributed and the support arrangements. For a beam of given length, the deflection per unit load (or stiffness) varies inversely as the product EI, which is called the "flexural rigidity." This quantity has significant implications for the selection of materials in deflection-limited design in bending.

There are close analogies between the analyses of bending and torsion (for regular axisymmetric shapes such as tubes and circular shafts). In this case, an applied torque generates shear stresses that increase linearly with radial distance from the axis of the member, and the response of the member is to twist. The stiffness is defined by the twist per unit length divided by the torque. The "torsional rigidity" is now the important parameter—the product of the shear modulus, G, and the polar second moment of area, J, defined by

$$J = \int r^2 \mathrm{d}A$$

Now it can be seen that efficient shapes are those which arrange the material at a large radius from the section axis, i.e., hollow tubes. Shape factors for torsion have also been defined in a similar way to those for bending (Ashby 1999).

2. *Lightweight Design in Bending*

A common design requirement is to resist deflection in bending while minimizing the component mass (e.g., in sports equipment such as golf clubs, or in lightweight transport such as bicycles). Here material selection requires a trade-off between Young's modulus, E (which governs stiffness) and density, ρ (which governs mass). The mass, m, of a beam of given length L and cross-sectional area A is simply

$$m = \rho L A$$

The deflection, however, depends on the second moment of area, which can vary widely at constant A (as indicated by the shape factors in Table 1). Material selection must therefore take account of geometric constraints imposed on the area and shape. This is the basis of the approach described by Ashby (1999) for optimizing the selection of materials using ''merit indices''. As a simple example, consider the selection of material to minimize the mass for a square beam of dimensions $B \times B$ and fixed length L, with a specified allowable deflection per unit load δ/W. The dimension B is a free variable. It will need to vary from one material to another in order to provide sufficient stiffness. The area $A(=B^2)$ must therefore be eliminated from the expression for the ''objective,'' mass, and this is achieved using an expression for the ''constraint,'' deflection per unit load. Noting that, for a square section, $I = B^4/12$, we proceed as follows:

$$m = \rho L B^2$$

$$\frac{\delta}{W} = C_1 \frac{12L^3}{EB^4}$$

so

$$B^2 = \left(\frac{12C_1L^3}{E(\delta/W)}\right)^{1/2}$$

Hence, the mass of the beam is

$$m = \left(\frac{12C_1L^5}{(\delta/W)}\right)^{1/2}\left(\frac{\rho}{E^{1/2}}\right)$$

All of the parameters in the first set of parentheses are specified by the design, so the mass only depends on the group of material properties in the second set of parentheses. To minimize mass one needs to maximize $E^{1/2}/\rho$—this quantity is called a ''merit index.'' An important point to note is that the result is not the ''intuitive'' answer, E/ρ, owing to the way mass and deflection depend independently on the cross-section geometry. Furthermore, if the constraints on geometry are modified, so is the merit index. For example, if the width of the beam is fixed but its depth is allowed to vary, the analysis yields the merit index, $E^{1/3}/\rho$. The significance of these results is best appreciated by illustrating the indices on a ''material selection chart.''

Figures 1 and 2 show a material selection chart—a plot of Young's modulus against density. Logarithmic scales are used since both properties span several orders of magnitude. Data for each class of material cluster into well-defined regions (reflecting the underlying atomic bonding and bond stiffness in each class).

On logarithmic scales, constant values of the merit indices plot graphically as straight lines, e.g., $E^{1/2}/\rho = constant$ is a line of slope 2. Guidelines of slope 1, 2, and 3 (representing the merit indices E/ρ, $E^{1/2}/\rho$, and $E^{1/3}/\rho$, respectively) are shown in Fig. 1. Materials lying along a given line all have equal mass for the specified design constraints (but different cross-sectional sizes according to their respective Young's moduli). To maximize a given index, the line is translated across the chart to leave a subset of materials, as shown in Fig. 2. Hence for maximum E/ρ, composites

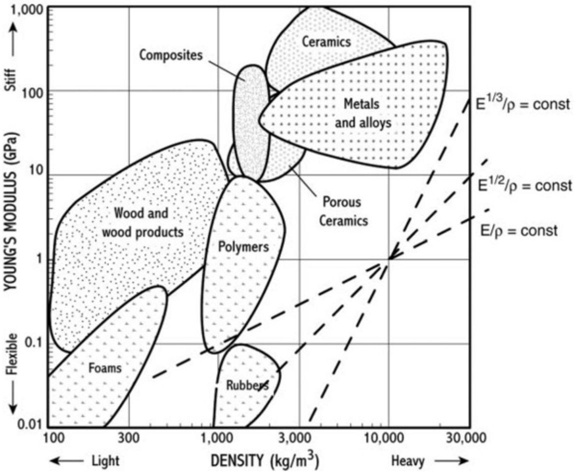


Figure 1
Material selection chart, showing Young's modulus against density for the material classes, with guidelines for three merit indices.

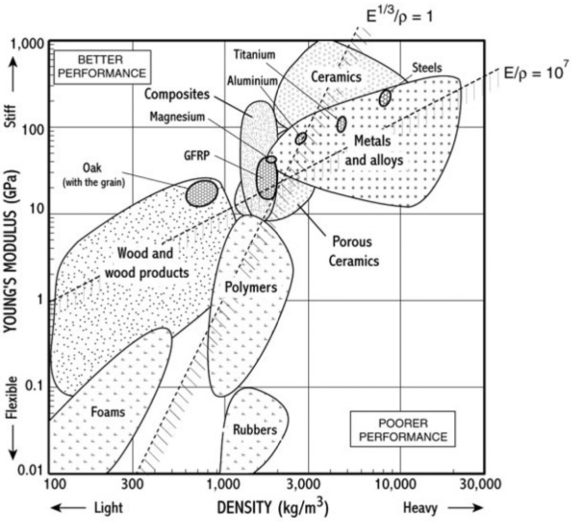


Figure 2
Material selection chart, showing Young's modulus against density for the material classes, with selected structural materials highlighted. Selection lines are shown for two merit indices.

and some metals compete, while ceramics are the best (although unlikely to be used on the grounds of poor toughness). The ranking of materials changes significantly for a different index; e.g., woods have a clear advantage over all metals for panels of fixed width ($E^{1/3}/\rho$). Composites and metals do, however, compete in this situation by exploiting sandwich panel construction (see below). Having identified candidate materials that satisfy the stiffness constraint at low mass, the choice will be further narrowed by imposing other design limitations (strength, cost, etc.).

2.1 Structural Efficiency

The analysis above assumes a constant cross-sectional shape for all materials (e.g., a square section, or a panel of fixed width). Where possible, designers will use sections with an efficient shape factor—but not all materials are equally easy to make into efficient shapes, owing to manufacturing constraints. Woods can only be made into solid sections, or thick-walled hollow box sections, while steels can be made into very efficient I-beams, and composites fall somewhere in between. Table 2 indicates typical limits on shape factor for various structural materials.

Selection of materials for lightweight design in bending must take account of whether it is an option to use shaping. A partial solution can be obtained by an extension to the merit index approach. In certain circumstances, the merit index for constant shape (e.g., $E^{1/2}/\rho$) may be modified to allow for shape by incorporating the shape factor in the index, e.g., $(\Phi E)^{1/2}/\rho$. The effect of shaping can be visualized as enhancing the modulus of the material, allowing the denser materials (such as steel) to become more competitive again on account of their large shape factor. Table 2 shows typical values for the merit index without shape ($E^{1/2}/\rho$) and with $(\Phi E)^{1/2}/\rho$. In the first case wood comes out on top, while the effect of shape is to put aluminum first, closely followed by glass fiber-reinforced polymer (GFRP).

As noted above, other constraints will come into play—strength, size, manufacturing problems, and cost. The cost of a structural section is influenced by more than the initial material cost (e.g., manufacturing route), but an indication of the effect of material cost can be obtained by a further modification to the merit index. The index $E^{1/2}/\rho$ is inversely proportional to the mass of the beam (at constant shape). Dividing this by the cost per unit weight of the material, C_m, gives a merit index which is inversely proportional to the material cost of the beam: $E^{1/2}/(\rho C_m)$. Table 2 shows typical values for this index—steel now stands out clearly above the rest, and would do even better if shape were included. This helps explain the widespread use of steel for structural design where cost dominates.

Composites are difficult to make into efficient structural shapes, but they can be used efficiently by making a sandwich panel. Sandwich panels use thin skins separated by a core of low Young's modulus and density, such as a polymer foam or honeycomb. Composites are easy to fabricate in thin sheets, and also to join to the core using adhesives. Sandwich panels are structurally efficient as the core material makes a negligible contribution to weight or stiffness, but the stiff, load-bearing material is held at maximum distance from the neutral axis, giving the best possible magnitude of I. The role of the core is simply to hold the skins apart. Examples of products exploiting sandwich panel construction are skis, aircraft floor panels, and helicopter rotors, using high Young's modulus composites (such as carbon fiber-reinforced polymer (CFRP)).

The core in a sandwich panel contributes virtually nothing to the overall flexural rigidity, but it must be stiff enough to prevent the panel skins from collapsing together. Cellular solids, such as foams and honeycombs are an important class of materials for providing modest stiffness (and strength) at very low weight. Apart from sandwich panel cores, these materials are exploited in a wide range of packaging materials. The properties of cellular solids are usually defined by their relative modulus (E/E_s) and relative density (ρ/ρ_s), i.e., the Young's modulus and density of the cellular material divided by the values for the solid cell wall. Typically these properties are related by an expression such as

$$\frac{E}{E_s} = \left(\frac{\rho}{\rho_s}\right)^n$$

Table 2
Relative performance of different classes of structural materials.

Material class	Typical maximum shape factor, Φ_{max}	$E^{1/2}/\rho$	$(\Phi_{max}E)^{1/2}/\rho$	$E^{1/2}/(\rho C_m)$
Wood (oak)	8	**4.1**	11.5	1.5
GFRP composite	39	2.9	19.0	1.0
6061 Al alloy	44	3.1	**20.5**	2.1
Mild steel	65	1.8	14.7	**4.5**

where $n \approx 1$–2. On the material selection chart, Fig. 1, cellular solids will therefore appear to the left and below the solid from which they are made. For example, woods are natural cellular solids made of polymeric fibers, and fall on the left of the chart at low Young's modulus. The porous structure is responsible for enhancing the structural efficiency of woods. There are many examples in nature of cellular structures giving structural efficiency, e.g., bones, leaves, and feathers.

3. *Future Challenges*

There is much research effort into mimicking the performance of cellular natural materials using engineering materials. One development area is the manufacture of metallic foams (mostly in aluminum) for use as lightweight core materials for automotive structures. Medical engineering is another growth area, where artificial implant materials are being developed to have properties much closer to those of bone.

Steady improvements are being made in manufacturing techniques in composite materials, allowing more efficient and complex shapes to be made, e.g., hollow bicycle frames which are also shaped to enhance aerodynamic efficiency. Developments in polymer fibers are also leading to competition with steels for heavy-duty cabling and ropes, e.g., tethers for oil rigs and cables for suspension bridges.

See also: Elastic Behavior of Cellular Solids*; Elastic Behavior of Composites*

Bibliography

Ashby M F 1999 *Materials Selection in Mechanical Design*, 2nd edn. Butterworth Heinemann, Oxford

Gibson L J, Ashby M F 1998 *Cellular Solids*, 2nd edn. Cambridge University Press, Cambridge

Lovatt A M, Shercliff H R, Withers P J 2000 *Material Selection and Processing*. Technology Enhancement Programme (TEP), London (further information: http://www-materials.eng.cam.ac.uk/mpsite/)

H. R. Shercliff and M. F. Ashby
University of Cambridge, UK

Elasticity in Wood and Concrete: Hygromechanical Effects

Many materials absorb moisture in high-humidity environments. In order to account for moisture effects in the design of wood and concrete structures, as well as advanced composite materials, it is necessary to know both the total moisture content and its profile as a function of the service environment. Knowledge of the diffusion characteristics of materials is also necessary for the purpose of performing accelerated environmental conditioning prior to mechanical loading. In addition to swelling or shrinkage, moisture adsorption and desorption can change the mechanical properties of materials. Both nonsymmetric moisture distribution and unsymmetric or antisymmetric material structure can cause hygromechanical warping. Therefore, understanding hygromechanical behavior is important not only for dimensional stability but also to ensure the safety of structures.

1. *Moisture in Cellular and Porous Materials*

Wood moisture may appear in four forms: (i) as water of constitution, (ii) as surface-bound water or water in an adsorbed monolayer with a high density ($1.3\,\mathrm{g\,cm^{-3}}$), (iii) as water adsorbed in multimolecular layers with a decreasing order of dipoles, and (iv) as capillary-condensed water. Constitutional water is due to the organic nature of cell walls. It cannot be removed without altering the chemical composition of the walls. Surface-bound water differs from the water of constitution and can be removed without altering the wood chemically. Since water is a polar liquid consisting of a negative hydroxyl $(\mathrm{OH})^-$ and a positive hydrogen fraction, it is assumed that hydroxyl groups are attached by valence bonds. Capillary-condensed water, as surface-bound water, can be freely evolved from wood.

The condition when all the cell-wall openings are filled with water, but no free water is present in the cell cavities, is the fiber saturation point (FSP). The movement of liquid water above the FSP is caused by capillary forces. Below the FSP bound water moves through the cell walls because of moisture gradients. The phenomenon is nevertheless complicated, because diffusion of moisture through the void structure, and bound water diffusion through the cell walls, occurs simultaneously (Kolmann and Côté 1968).

As in wood, the water contained in cement paste and concrete can be classified into types depending on the degree of difficulty of its removal. Chemically bound water is combined with the fresh cement in the hydration reaction and is not lost on drying but is only evaporated when the paste is decomposed by heating to high temperatures in excess of 1000 °C. Interlayer water is displaced in gel pores and can be removed only by strong drying. Adsorbed water is the water that lies close to solid surfaces under the influence of surface attractive forces. A large proportion of this water can be lost on drying to 30% relative humidity. Capillary water is located in the capillary and larger gel pores. The larger voids may only be partially filled with water but the remaining space contains water vapor. The movement of moisture in concrete is caused by capillary forces and moisture gradients and depends on the rate of hydration reactions with cement.

Moisture diffusion models are based on Fick's law. Available data (Skaar 1988) indicate classical Fickian behavior in terms of total weight gains up to an apparent equilibrium level for both homogeneous materials and composites.

2. *Hygrodeformation of Materials*

2.1 *Swelling and Shrinkage of Wood and Wooden Materials*

In the living tree, the cell walls of wood are always in the fully swollen condition, at fiber saturation. When green wood dries, hygroscopic shrinkage takes place, the extent depending on a number of factors. These include the amount of moisture loss, the structural direction (tangential, radial, or longitudinal), the kind of wood, drying stresses caused by moisture gradients, etc. Wood is normally dried to a moisture content approximately equal to its anticipated mean equilibrium moisture content (EMC) in use. However, since environmental conditions are rarely constant the EMC is continually changing and the moisture content of wood is seldom equal to its anticipated EMC.

Hygroexpansion is given in terms of the percentage shrinkage or swelling. It depends on the deformation of wood itself and the fraction of air spaces in addition to cell-wall material. Hygrodeformation is not linear with moisture changes (Kolmann and Côté 1968).

The cell walls of wood are anisotropic, and therefore hygromechanical properties are also anisotropic. Typically, longitudinal shrinkage is least, usually 0.1–0.3% from the FSP to the completely dry condition. Total radial shrinkage ranges from 3% to 6% and tangential shrinkage from 6% to 12% for most woods over the same moisture range. This anisotropy is responsible for much of the warping of boards cut from various locations in the cross-section of a log.

In the production of wooden composites (plywood, flakeboards, fiberboards, etc.), plastic deformation is the major component of wood deformation that occurs during the hot-pressing process. When moistening, plastic deformation becomes recoverable and there is a large difference between swelling of the densified wood in the direction of pressing and in the transverse direction. The reconstituted wood structure means that the specific moisture strain of the densified wood in comparison with normal wood increases nonlinearly, when moisture content grows, in the direction of pressing but decreases in the transverse direction.

The methods of achieving dimensional stabilization of wood and wooden composites are: (i) cross-laminating of anisotropic sheets to restrain the dimensional changes; (ii) applying water-resistant surface and internal coatings to retard moisture adsorption or loss; (iii) reducing the hygroscopicity of the wood by heat treatment; (iv) introducing a nonvolatile material into the cell walls thus reducing their capacity for water uptake; and (v) cross-linking of the cellulose chains of the fibers by resin impregnation, acetylation, and treatment with polystyrene.

2.2 *Shrinkage and Swelling of Concrete*

There are two causes of shrinkage: drying and carbonation. Water can only be lost from the surface and therefore the inner core of a specimen constrains overall movement and hence the measured shrinkage varies with sample size. Maximum shrinkage occurs on the first drying, and a considerable part of this is irreversible, i.e., it is not recovered on rewetting (Illston 1998). During immersion in water, the water drawn into the material causes a net increase in volume. This is a characteristic of many gels, but in concrete the expansion is resisted by the skeletal structure so that the swelling is small compared to the drying shrinkage strains.

Four principal mechanisms have been proposed for shrinkage and swelling in cement pastes and concrete: capillary tension, surface tension, disjoining pressure, and movement of interlayer water. The drying shrinkage of concrete is less than that of neat cement paste because of the restraining influence of the aggregate. Apparently concrete has a shrinkage of 10–30% of that of the paste. A member with a large surface area to volume ratio will dry and therefore shrink more rapidly than a beam with a square cross-section of the same area. Carbonation shrinkage differs from drying shrinkage in that its cause is chemical, and it does not result from loss of water.

To minimize shrinkage of concrete it is desirable to: (i) keep the amount of mixing water to a minimum; (ii) cure the concrete well; (iii) use construction joints to control the position of cracks; (iv) use shrinkage reinforcement; and (v) use appropriate dense and nonporous aggregates.

2.3 *Effect of Moisture on Advanced Composites*

Fiber-reinforced epoxy laminates and other advanced composites absorb atmospheric moisture, apparently by instantaneous surface adsorption and subsequent diffusion through the interior. The absorbed water is not liquid, but exists as hydrogen-bonded molecules or clusters within the polymer. Liquid water may, however, be transported by capillary action along cracks and along fiber–matrix interfaces and may appear at interior voids. The absorbed water softens epoxy resins, and causes them to swell and deform (Gillat and Broutman 1978). The structural consequences of these effects are not yet fully quantified.

2.4 Hygroexpansion and Mechanical Stress

Materials swell to an extent that depends on the environmental conditions. One should expect, therefore, that elements restrained from swelling should exert a stress in the direction in which they are restrained. The swelling pressure of the material is limited by the crushing strength, which is a function of its density, moisture content, and temperature. The swelling pressure reaches a peak and then decreases with time. This is caused by simultaneous occurrence of swelling and stress relaxation with time.

The anisotropy of wood is responsible for the warping of lumber associated with initial drying and during subsequent use. The most important practical effects are different kinds of transverse warping caused by tangential/radial anisotropy in shrinkage. This effect can be reduced if the lumber is restrained from cupping and warping during drying by the application of weights or other restraints. However, hygrodeformation cannot be eliminated completely since some residual stresses are present.

3. Modeling of Hygromechanical Effects

For the estimation of mechanical properties, stresses, and deformation of a material under environmental effects, models are useful. In one approach, the material is divided into a system of elementary layers (real or conditioned) (Brauns and Rocens 1997). The hygromechanical properties of the layers are either determined from the properties of the components using analytical methods or found experimentally. In the general case, the stresses in elementary layers are different and the stress state of the material is inhomogeneous. By using a static equivalent system of average force stresses and couple stresses acting on a unit cube of material, the midplane strains and the curvatures depending on relative moisture content can be determined.

The force and couple stresses in the unit volume of material are calculated by averaging the stresses in the elementary layers of the material. Stresses in the layers are determined using specific environmental strains as well as the stiffness of the material determined analytically. The results calculated can be compared with values found experimentally of swelling deformation and curvatures (Fig. 1). The moisture profile perpendicular to the plane of layers is nonlinear, and the difference of moisture content of layers after seven days' exposure is 9.0%.

4. Concluding Remarks

Hygroexpansion and warping depend on the moisture content and profile within the material. Hygromechanical effects should be determined based on the configuration of the composite, precise stacking sequence of the layers, as well as the ply stiffnesses. The optimum configuration for hygroeffects may not coincide with that for mechanical properties, and for every type of loading, specific composite structure should be used. Special methods of treatment have to be used for dimensional stabilization of wood and wooden composites as well as concrete used in structures.

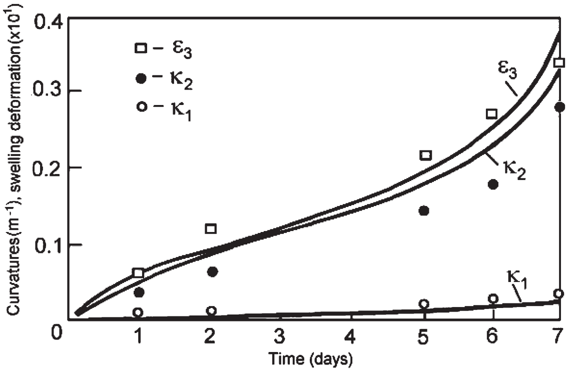


Figure 1
Experimental and calculated values of curvatures (κ_i) and swelling deformations (ε_3) perpendicular to the plane of orthotropic wood laminate as a function of duration of unilateral exposure to moisture (after Rocens 1983).

Bibliography

Brauns J, Rocens K 1997 Hygromechanical behaviour of wooden composites. *Wood Sci. Technol.* **31**, 193–204

Gillot O, Broutman L J 1978 Effect of an external stress on moisture diffusion and degradation in a graphite-reinforced epoxy laminate. In: Vinson J R (ed.) *Advanced Composite Materials–Environmental Effects*. American Society for Testing and Materials, Philadelphia,, pp. 61–83

Illston J M 1998 *Construction Materials: Their Nature and Behaviour*. Spon, New York

Kolmann F F P, Côté W A 1968 *Principles of Wood Science and Technology*. Springer, Berlin

Rocens K 1983 Macrostructure theory of modification of wood properties. *J. Appl. Polym. Sci.* **37**, 923–41

Skaar C 1988 *Wood–Water Relations*. Springer, New York

J. Brauns[a] and K. Rocens[b]
[a]*Latvia University of Agriculture, Jelgava, Latvia*
[b]*Riga Technical University, Latvia*

Elastomers, Ceramic-modified

Many elastomers are generally compounded with a permanent reinforcing filler, particularly those which cannot undergo strain-induced crystallization (Medalia and Kraus 1994, Warrick *et al.* 1979). The two most important examples are the addition of carbon black to natural rubber and to some synthetic elastomers, and the addition of silica to siloxane rubbers (Warrick *et al.* 1979). In fact, the reinforcement of natural rubber and related elastomeric materials is one of the most important processes in elastomer technology. Not surprisingly, there is an incredible amount of relevant experimental data available, with most of these data relating to reinforcement of natural rubber by carbon black (Donnet and Vidal 1986). However, other polymers such as poly(dimethylsiloxane), and other fillers such as precipitated silica, clays, metallic particles, and even glassy polymers, have become of interest (Calvert 1996, Giannelis 1996, Mark 1995, Mark 1996a, Mark 1996b, Mark *et al.* 1995, Schmidt 1996, Wen and Wilkes 1996). Some of these fillers can now be generated *in situ*, within elastomeric polymers, and this the main subject of the present article.

1. *General Comments on* In Situ *Reinforcement*

The polymer which has been used most extensively in studies of *in situ*-precipitated fillers is poly(dimethylsiloxane) (PDMS) $[–Si(CH_3)_2O–]_x$. It has been chosen in part because it is compatible with many of the organometallic materials used to generate ceramic-type phases, and also because it is an elastomer which requires considerable reinforcement from silica or some other filler before it is useful in most industrial applications (Warrick *et al.* 1979).

The idea of *in situ* generation of reinforcing fillers in polymers is based on an important advance in the ceramics area. It involves the use of preparative techniques that are heavily based on chemical reactions—for example the generation of ceramic-type materials by the hydrolysis of an organometallic compound (Cheetham *et al.* 1994, Mark *et al.* 1995). In the ceramics area, the advantages are the purity of the resulting products, the possibility of avoiding high temperatures (thereby permitting survival of organic guest molecules), the control of ultrastructure (at the nanometer level), the relative ease of forming ceramic "alloys," and the possibility of incorporating a different material into the resulting ceramic while it is still porous, prior to its densification. In the present application to elastomers, the new approach has the advantages of avoiding the difficult, time-consuming, and energy-intensive process of blending agglomerated filler into high-molecular-weight polymers, and simplifying the problem of obtaining good dispersions. The main disadvantage is the increased cost, and the reluctance of industry to carry out the major retooling required to introduce such a different technology. In any case, this approach has been investigated using a variety of fillers, for example silica (SiO_2) by hydrolysis of organosilicates, titania (TiO_2) from titanates, alumina (Al_2O_3) from aluminates, zirconia (ZrO_2) from zirconates, etc. (Clarson and Mark 1993, Schaefer and Mark 1990).

The following sections consist mainly of descriptions of how some of these sol–gel ideas can be extended using concepts from the area of polymer chemistry (Mark 1996b). The modifications are generally carried out by using elastomeric polymer chains terminated with functional groups that can participate in the ceramic-generating reaction, typically hydroxy groups in the case of a hydrolysis reaction. This permits intimate bonding between the inorganic (ceramic) phase and the organic (polymer) phase in a way that is most likely to give rise to novel mechanical properties.

Although the reinforcement of elastomers by dispersed ceramic phases is of primary interest in the present context, the case where the ceramic predominates and becomes the continuous phase (Mark 1996b) is also mentioned. The elastomeric phase therefore becomes dispersed in it, improving its mechanical properties, such as impact resistance. Yet another approach to the novel reinforcement of elastomers involves some related systems in which there are either magnetically responsive particles or crystalline zeolites (Erman and Mark 1997).

2. *Preparation*

The most important reaction of this type involves the catalytic hydrolysis of tetraethoxysilane, also known as tetraethylorthosilicate (TEOS) (Erman and Mark 1997, Mark 1992, Mark *et al.* 1993):

$$Si(OC_2H_5)_4 + 2H_2O \rightarrow SiO_2 + 4C_2H_5OH \qquad (1)$$

In the original sol–gel technique, the gel thus formed is first dried to remove unreacted TEOS, ethanol, water, and catalyst (which is generally chosen to be volatile). It is then fired into a porous ceramic, which may then be densified into the final ceramic object.

In the simplest application of this approach to obtaining elastomer reinforcement, some of the organometallic material is absorbed into the cross-linked network, and the swollen sample placed into water containing the catalyst (typically a volatile base such as ammonia or ethylamine). Hydrolysis to form the desired silica-like particles proceeds rapidly at room temperature to yield the order of 50 wt.% filler in less than an hour (Erman and Mark 1997). Alternatively, hydroxy-terminated chains are blended with enough TEOS both to end link them and to provide silica by the hydrolysis reaction. Thus, curing and filling take place simultaneously, in a one-step process. In an alternative method, TEOS is blended into a polymer having end groups (e.g., vinyls) that are unreactive under the hydrolysis conditions. The silica is then formed in the usual manner (Eqn. (1)) and the product dried. The resulting slurry of polymer and silica is stable and can be cross-linked later using any of the standard cross-linking techniques, such as vinyl–silane coupling or peroxide thermolysis (Erman and Mark 1997).

3. Electron Microscopy

Both transmission and scanning electron microscopy have been used to characterize these novel composite materials (Mark 1989). The information obtained in this way includes: (i) the nature of the precipitated phase (particulate or nonparticulate); (ii) the average particle size, if particulate; (iii) the distribution of particle sizes; (iv) the degree to which the particles are well defined (smoothness of the interfaces); and (v) the degree of agglomeration of the particles (Mark 1996b).

Such results show that the particles formed in this manner typically have a narrow distribution of sizes, with most diameters falling in the 200–300 Å range (Erman and Mark 1997, Xu *et al.* 1990). They are generally well dispersed and essentially unagglomerated, which suggests that the generation reaction may involve simple homogeneous nucleation. This is consistent with the fact that particles growing independently of one another and separated by cross-linked polymer would not agglomerate unless very high concentrations were reached. There is some evidence that at least in some cases, the pore size of the network constrains the growth of these particles, in that the average diameters observed seem to increase with decrease in cross-link density (Breiner and Mark 1998, Ulibarri *et al.* 1992).

One interesting result from characterization studies of this type is the conclusion that basic catalysts generally yield particles that are well defined, whereas acidic catalysts yield particles that are rather "fuzzy" (Erman and Mark 1997, Mark *et al.* 1985).

4. Scattering Techniques

A number of x-ray and neutron scattering studies has been carried out on these filled elastomers (Landry *et al.* 1995, Mark 1996b, McCarthy *et al.* 1998). In general, such data can provide estimates of average particle size and possibly particle-size distribution. In addition, the terminal slopes give an indication of the nature of the interfaces, with -3 corresponding to rough interfaces and -4 to smooth.

5. Nuclear Magnetic Resonance

Another way to characterize the *in situ*-generated particles is by nuclear magnetic resonance (NMR) imaging, using ^{1}H and ^{29}Si magic-angle spinning, with two-dimensional Fourier transform spin-echo techniques (Erman and Mark 1997, Mark 1996a). The result of greatest interest is the distribution of particles within the elastomer, and how the precipitation reaction front moves into the sample with time. The method is obviously nondestructive, but slices taken from the sample can be further studied by density measurements in a gradient column, by electron microscopy, or by x-ray or neutron scattering. Elastomer which has been infiltrated into a silica aerogel can also be imaged in this way (Erman and Mark 1997).

6. Aging

Permitting precipitated silica particles to remain in contact with their aqueous catalyst solution can permit them to "age" or "digest" (Mark 1996b, Xu *et al.* 1990). Electron microscopy results suggest that some reorganization is occurring, with the particles becoming better defined, more uniform in size, and possibly even less aggregated. There seem to be interesting parallels with "Ostwald ripening" in the area of colloid science.

7. Densities

Comparisons between the values of wt.% filler obtained from density measurements and the values obtained directly from weight increases can yield very useful information on the filler particles. For example, the fact that the former estimate is smaller than the latter in the case of silica-filled PDMS elastomers (Erman and Mark 1997, Ulibarri *et al.* 1992) indicates that there are probably either voids or unreacted organic groups in the filler particles, and possibly both. Incomplete conversion of TEOS to silica is also indicated by ^{29}Si NMR measurements (Ulibarri *et al.* 1992).

8. Calorimetry

Differential scanning calorimetry measurements at low temperatures have been carried out on PDMS elastomers containing *in situ*-precipitated silica (Mark 1996a). The presence of the silica was found to reduce both the extent of crystallization and the rate of crystallization when the elastomers were in the unstretched state. This is in interesting contrast to similar studies of PDMS in the stretched state, where the filler may facilitate the crystallization process.

9. Thermogravimetric Analysis

In at least some cases, *in situ*-precipitated fillers may have the advantage of increasing thermal stability (Erman and Mark 1997). A possible mechanism for this improvement would be increased capability of the *in situ*-produced silica to tie up hydroxy chain ends that participate in the degradation reaction.

10. Mechanical Properties and Equilibrium Swelling

10.1 Uniaxial Extension

The reinforcing ability of such *in situ*-generated particles has been amply demonstrated for a variety of deformations, including uniaxial extension (simple elongation), biaxial extension, shear, and torsion (Erman and Mark 1997, Mark 1996b, Mark and Erman 1988). Illustrative results obtained for elongation (reciprocal elongation $\alpha^{-1} < 1$) and compression ($\alpha^{-1} > 1$) are depicted schematically in Fig. 1, where the modulus is plotted against α^{-1}. In the case of uniaxial extension, the modulus [f*] frequently increases by more than an order of magnitude, with the isotherms generally showing the upturns at high elongation that are the signature of good reinforcement (Erman and Mark 1997, Mark *et al.* 1992). As is generally the case in filled elastomers, there is considerable irreversibility in the isotherms, which is thought to be due to irrecoverable sliding of the chains over the surfaces of the filler particles. Some fillers other than silica, for example titania, do give stress–strain isotherms that are reversible, indicating interesting differences in surface chemistry (Mark 1996b).

10.2 Adhesion Between Phases

It is possible to interpret equilibrium swelling measurements obtained on unfilled and filled elastomers to estimate the degree of adhesion between elastomer and filler particles (Kraus 1965, Mark 1996b). Some typical results for *in situ*-generated fillers are presented schematically in Fig. 2, in which the ordinate is the ratio of the extent of swelling of a filled network relative to its swelling in the unfilled state. The results differ greatly from those expected for nonadhering fillers, indicating good bonding between the two phases. Resistance to separation from the surface in such swelling tests, however, does not necessarily contradict the conclusion that the chains have mobility along the surface, as was described in Sect. 10.1.

10.3 Biaxial Extension

Biaxial extension is one of the most important of the mechanical deformations other than simple elongation, and can be imposed by inflation of a sheet of the elastomer. Since biaxial extension is equivalent to uniaxial compression, both types of data can be combined in the same plot, as shown in Fig. 1. Very strong reinforcing effects from the precipitated silica occur for biaxial extension as well as for elongation. This is evident from both the large upward shifts of the isotherms as a whole and from the pronounced upturns at both high elongations ($\alpha^{-1} < 1$), and at high compressions (biaxial extensions) ($\alpha^{-1} > 1$). The pronounced maxima and minima in the isotherms for the filled elastomers in biaxial extension are not understood at present (Erman and Mark 1997), and

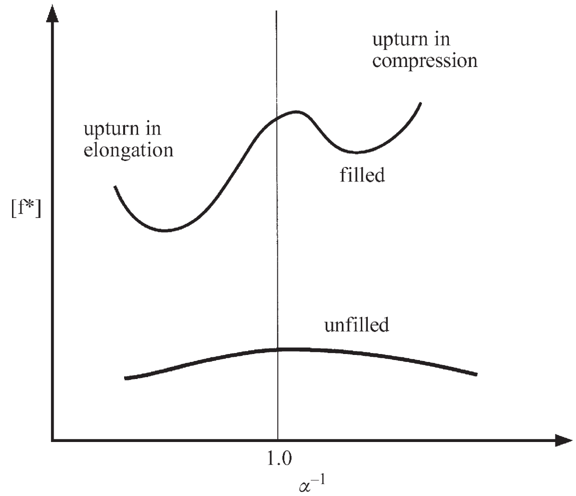


Figure 1
Schematic stress-strain isotherms for PDMS-silica in-situ reinforced elastomers in elongation (region to the left of the vertical dashed line, with reciprocal elongation $\alpha^{-1} < 1$), and in biaxial extension (compression) (to the right, with $\alpha^{-1} > 1$).

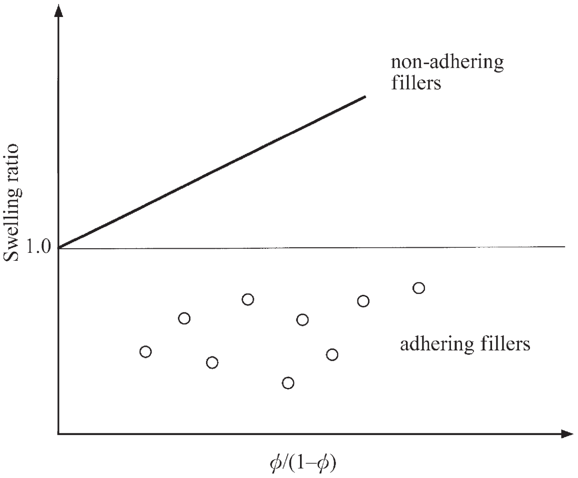


Figure 2
Plot of swelling ratio characterizing the swelling of a filled PDMS network relative to that of an unfilled PDMS network, against filler loading expressed as volume ratio of filler to rubber $\phi/(1-\phi)$ (where ϕ is the volume fraction of filler). The symbols represent experimental results obtained using silica, titania, or silica–titania mixed-oxide fillers.

represent challenges to theories on the origin of reinforcing effects in elastomers in general.

10.4 Shear

Pure shear is conveniently imposed by stretching a wide rectangular sheet in a direction perpendicular to its width. As in the case of uniaxial extension and compression, the addition of filler shifts the isotherms upwards to significantly higher values of the modulus. In the case of the filled networks there is an initial decrease in the modulus with increase in deformation. This may be due to stress-induced rearrangements of the chains in the vicinity of the filler particles (Erman and Mark 1997).

The filled samples have much higher shear moduli than the corresponding unfilled samples. Also, at high deformation, the filled samples show pronounced upturns in the reduced stress or modulus, which indicate very good reinforcement. The magnitudes of the upturns are much less pronounced, however, than those in uniaxial extension or compression.

10.5 Torsion

As was the case for the other deformations, an increase in filler content was found to shift the torsion isotherms upward, to significantly higher values of the shear modulus (Erman and Mark 1997). These curves are significantly different from the ones for the other types of deformation, however, in that there is little evidence for upturns in the modulus at high deformations. This could be due either to the nature of the deformation or to the inability, to date, of reaching sufficiently high torsional deformations.

10.6 Hardness

Another illustrative property of importance in such hybrid organic–inorganic composites is hardness. The case of the PDMS–SiO_2 system is illustrated in Fig. 3 (Mark 1996a). The hardness can be varied greatly by changing the organic-to-inorganic molar ratio, as measured by the molar ratio of organic R groups (in this case methyl) to silicon atoms. Low values of the R:Si ratio yield a brittle ceramic, and high values yield an *in situ*-reinforced elastomer. The most interesting range of values, where R:Si $\sim$ 1, can give a hybrid material that can be viewed either as a ceramic of reduced brittleness or an elastomer of increased hardness, depending on one's point of view.

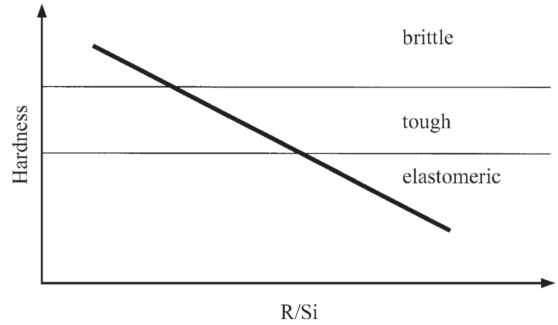


Figure 3
Dependence of the D-scale hardness of PDMS composites on the molar ratio of alkyl groups to silicon atoms.

10.7 Dynamic Mechanical Properties

Measurement of dynamic mechanical losses can provide important information on transitions occurring in polymers, including filled elastomers. Such results, for example, document the effect of filler on an elastomer's glass transition temperature and melting point (Mark 1996b).

10.8 Cyclic Deformations

In a typical experiment of this type, silica-filled PDMS elastomers are subjected to cyclic compressions, with the modulus and sample dimensions monitored as a function of the number of cycles, the time period of cycling, or the power consumed (Mark 1996b). Of particular interest in such experiments is the decrease in fractional length of a sample with time, since this characterizes the amount of creep or compression set resulting from the cyclic deformation. Unfilled PDMS samples and those filled by blending with fume silica show considerable compression set, while the samples containing *in situ*-precipitated silica show relatively little (Erman and Mark 1997). This suggests that the *in situ* silica can retard the chemical changes caused by cyclic stresses as well as those caused by high temperatures, and is thus consistent with the thermogravimetric analysis results already cited (Mark 1996b).

10.9 Impact Resistance

Impact strengths of PDMS–SiO_2 samples have been determined by the Charpy pendulum impact test and by the falling-weight impact test (Erman and Mark 1997). The samples investigated were PDMS-modified SiO_2 and SiO_2/TiO_2 glasses with PDMS contents ranging from 0 to 65 wt.%. As expected, only samples with relatively high ceramic contents were sufficiently brittle to be studied in this manner. It was reported that the larger the amount of PDMS used, or the higher its molecular weight, the higher the impact strength of the composite. Thus, considerable toughening of the glass can be achieved by increasing the amount of PDMS introduced. The effect of increase in PDMS molecular weight may be to increase the phase separation which leads to the energy-absorbing elastomeric domains (Erman and Mark 1997).

Microscopic observation of fracture surfaces can also provide useful information for interpreting impact strengths in terms of the physical processes involved. Specifically, smooth fracture surfaces are associated with completely brittle failure, with little effective resistance to either initiation or propagation of cracks. The samples with high PDMS contents had fracture surfaces showing some degree of "whitening" or shearing. This suggests a ductile, energy-absorbing response to the impacts, with increased resistance to crack propagation. Materials of this type are discussed in detail in the articles *Polymer-modified Ceramics** and *Polymer–Inorganic Composites, Properties of*#.

10.10 Equilibrium Swelling

In a swelling experiment, the network is typically placed into an excess of solvent, which it absorbs until the dilational stretching of the chains prevents further absorption of solvent (Erman and Mark 1989, Mark and Erman 1988). Adsorption of some network chains on to the particles of a reinforced elastomer would, of course, reduce the extent of swelling at equilibrium (as described above) and lead to inhomogeneous strain fields. Characterization of such equilibrium swelling as a function of filler constitution, filler amount, and the method of generating and incorporating it into an elastomer are the issues of greatest interest with regard to this type of deformation.

11. Comparisons Among Various Silica-based Fillers

As is now obvious, there is a variety of ways to generate silica-type fillers useful for reinforcing PDMS networks, and the extent to which such fillers provide reinforcement has been characterized (Mark 1996a). The best mechanical properties seem to be obtained either by precipitating silica directly into a swollen PDMS network after curing, or by incorporating silica which has been prepared from partially hydrolyzed TEOS with incorporation of some PDMS. A better molecular understanding of filler reinforcement in general could provide a detailed interpretation of these results.

12. Other Elastomers

Most of the studies to date on elastomer–ceramic materials have involved PDMS, primarily because of its great miscibility with TEOS. Similar studies (Mark and Erman 1988) on the related polymer, poly(methylphenylsiloxane), however, are also of considerable importance because of the stereochemically irregular structure of this polymer which prevents it from undergoing strain-induced crystallization (Erman and Mark 1997, Mark and Erman 1988). Good *in situ*-generated reinforcement was also achieved in this polymer, demonstrating that such crystallization is not important for this type of reinforcement (Mark 1996a).

A large number of other polymers has been treated in this way (Mark 1996a, Mark 1996b). For example, the same techniques have been shown to give good reinforcement in polyisobutylene elastomers and in poly(ethyl acrylate). In the case of some hydrocarbon polymers, miscibility with the alkoxysilane $Si(OR)_4$ can be improved by choosing R groups that will make the molecule more hydrophobic than TEOS.

13. Other Ceramic-type Fillers

Silica particles in PDMS elastomers can be a problem at high temperature, since the silanol groups on their surfaces can cause degradation of the polymer (Warrick *et al.* 1979). Because of this, and for other reasons, a variety of other fillers has been precipitated into PDMS and other elastomers (Mark 1996a, Mark 1996b), including titania, alumina, and zirconia, and these nonsilica fillers also provide good reinforcement. One interesting difference, however, is the already mentioned observation that the stress–strain isotherms in these cases frequently have much better reversibility (Erman and Mark 1997, Mark and Erman 1988). The differences are presumably due to different interactions between the groups present on these particle surfaces and the PDMS elastomeric matrix.

There is now increasing interest in introducing fillers consisting of more than one oxide (Mark 1996a, Mark 1996b). One of the hopes is that there will be an advantageous "delegation of responsibilities" in these mixed oxides, with one type of particle improving one set of properties and the other improving a different set.

14. Preparation of Bicontinuous Systems

At some compositions and under some hydrolysis conditions, bicontinuous phases can be obtained (with the silica and polymer phases interpenetrating one another). Some of the evidence for such interpenetration is from small-angle x-ray scattering and small-angle neutron scattering profiles (Mark 1996b).

Other evidence is from stress–strain isotherms in elongation. Since the two networks interpenetrate one another, the mechanical properties exhibited by the material can be very peculiar. In the first deformation of the virgin material, the silica network can give a very high initial modulus, but once this structure is broken, additional deformation cycles can indicate much lower values of the modulus. The mechanism for the generation of bicontinuous materials may be spinodal decomposition, occurring either before or after polymerization.

Bibliography

Breiner J M, Mark J E 1998 Preparation, structure, growth mechanisms and properties of siloxane composites containing silica, titania or mixed silica–titania phases. *Polymer* **39**, 5483–93

Calvert P 1996 Biomimetic inorganic–organic composites 315–36. In: Mann S (ed.) *Biomimetic Materials Chemistry*. VCH, New York

Cheetham A K, Brinker C J, Mecartney M L, Sanchez C (eds.) 1994 *Better Ceramics Through Chemistry VI*. Materials Research Society, Pittsburgh, PA

Clarson S J, Mark J E 1993 Siloxane elastomers . Clarson S J, Semlyen J A (eds.) *Siloxane Polymers*. Prentice Hall,. Englewood Cliffs, NJ, pp. 616–48

Donnet J-B, Vidal A 1986 Carbon black: surface properties and interactions with elastomers. *Adv. Polym. Sci.* **76**, 103–27

Erman B, Mark J E 1989 Rubber-like elasticity. *Annu. Rev. Phys. Chem.* **40**, 351–74

Erman B, Mark J E 1997 *Structures and Properties of Rubberlike Networks*. Oxford University Press, New York

Giannelis E P 1996 Organoceramic nanocomposites. In: Mann S (ed.) *Biomimetic Materials Chemistry*. VCH Publishers, New York, pp. 337–59

Kraus G (ed.) 1965 *Reinforcement of Elastomers*. Interscience, New York

Landry M R, Coltrain B K, Landry C J T, O'Reilly J M 1995 Structural models for homogeneous organic–inorganic hybrid materials: simulations of small-angle x-ray scattering profiles. *J. Polym. Sci., Polym. Phys.* **33**, 637–55

Mark J E 1989 Generate reinforcing particles in place. *CHEMTECH* **19**, 230–3

Mark J E 1992 Novel reinforcement techniques for elastomers. *J. Appl. Polym. Sci., Appl. Polym. Symp.* **50**, 273–82

Mark J E 1995 Some organic–inorganic nanocomposites. In: Martinez R J, Arris H, Emerson J A, Pike G (eds.) *Diversity into the Next Century*. SAMPE. Covina, CA, pp. 1–10

Mark J E 1996a Ceramic-reinforced polymers and polymer-modified ceramics. *Polym. Eng. Sci.* **36**, 2905–20

Mark J E 1996b The sol–gel route to inorganic–organic composites. *Hetero. Chem. Rev.* **3**, 307–26

Mark J E, Eisenberg A, Graessley W W, Mandelkern L, Samulski E T, Koenig J L, Wignall G D 1993b *Physical Properties of Polymers*. American Chemical Society, Washington, DC

Mark J E, Erman B 1988 *Rubberlike Elasticity. A Molecular Primer*. Wiley-Interscience, New York

Mark J E, Lee C Y-C, Bianconi P A (eds.) 1995 *Hybrid Organic–Inorganic Composites*. American Chemical Society, Washington

Mark J E, Ning Y-P, Jiang C-Y, Tang M-Y, Roth W C 1985 Electron microscopy of elastomers containing in situ-precipitated silica. *Polymer* **26**, 2069–72

Mark J E, Wang S, Xu P, Wen J 1992 Reinforcement from *in situ*-precipitated silica in polysiloxane elastomers under various types of deformation. In: Baney, R H, Gilliom, L R, Hirano, S-I, Schmidt, H K (eds.) *Submicron Multiphase Materials*. Materials Research Society, Pittsburgh, PA

McCarthy D W, Mark J E, Schaefer D W 1998 Synthesis, structure, and properties of hybrid organic–inorganic composites based on polysiloxanes. I. Poly(dimethylsiloxane) elastomers containing silica. *J. Polym. Sci., Polym. Phys.* **36**, 1167–89

Medalia A I, Kraus G 1994 Reinforcement of elastomers by particulate fillers. In: Mark J E, Erman B, Eirich F R (eds.) *Science and Technology of Rubber*. Academic Press, New York, pp. 378–418

Schaefer D W, Mark J E (eds.) 1990 *Polymer-based Molecular Composites*. Materials Research Society. Pittsburgh, PA

Schmidt H K 1996 Synthesis and applications of inorganic/organic composite materials. *Macromol. Symp.* **101**, 333–42

Ulibarri T A, Beaucage G, Schaefer D W, Olivier B J, Assink R A 1992 Molecular weight dependence of domain structure in silica–siloxane molecular composites. In: Baney R H, Gilliom L R, Hirano S-I, Schmidt H K (eds.) *Submicron Multiphase Materials*. Materials Research Society, Pittsburgh, PA, pp. 85–90

Warrick E L, Pierce O R, Polmanteer K E, Saam J C 1979 Silicone elastomer developments 1967–1977. *Rubber Chem. Technol.* **52**, 437–525

Wen J, Wilkes G L 1996 Organic/inorganic composites. In: Salamone J C (ed.) *Polymeric Materials Encyclopedia: Synthesis, Properties, and Applications*. CRC Press, Boca Raton, FL, pp. 4782–92

Xu P, Wang S, Mark J E 1990 Particle growth and development during the *in situ* precipitation of silica in a polymeric matrix . Zelinski B J J, Brinker C J, Clark D E, Ulrich D R (eds.) *Better Ceramics Through Chemistry IV*. Materials Research Society, Pittsburgh, PA, pp. 445–51

J. H. Mark
University of Cincinnati, Ohio, USA

Elastomers, Reinforcement of

Finely divided particulate fillers have been used in the rubber industry for centuries, as cheap diluents and to increase the modulus of elasticity. Reinforcement, signifying an increase in tensile strength, abrasion resistance, and other ultimate (i.e., failure) properties as well as modulus, was reported in 1905 (with zinc oxide). Soon thereafter the much stronger reinforcement by carbon black was found with tires (see *Rubber Tires**), which at that time were based entirely on natural rubber. Noncrystallizing elastomers such as SBR are highly dependent on reinforcement, and the synthetic rubber industry of today could not exist without the use of reinforcing agents.

1. Reinforcing Fillers

Carbon black is the pre-eminent reinforcing filler. More than 30 grades of carbon black are manufactured for use in rubber; these impart a broad spectrum of properties to the rubber compound. Since carbon black is discussed *in extenso* elsewhere (see *Networks, Failure of*[#]) only a brief review of the nature and important characteristics of carbon black is given here.

Carbon black consists of solid carbon entities of colloidal dimensions ($<1\,\mu$m) called *aggregates*. Each aggregate is composed of many particles fused together in a randomly arranged clustered (aciniform) morphology. The specific surface area is inversely proportional to the particle size (diameter), while the number and arrangement of particles in the aggregates determine the bulkiness, or ratio of the effective volume to the volume of carbon, which for historical reasons is called *structure*. Surface area is measured by adsorption of a gas (nitrogen) or of a surfactant (CTAB) in aqueous solution; adsorption of iodine is a convenient, though less accurate, method of measurement. Structure is measured by adsorption of a liquid (dibutyl phthalate), either on the carbon black as received (DBPA) or after a crushing and sieving treatment (CDBP). The surface of carbon black is composed of somewhat imperfect graphitic layers; these appear to be organized in a shingled arrangement, with exposed edges at which the carbon atoms may be bonded to hydrogen atoms or to oxygen groups, predominantly phenolic. The fraction of the surface covered by oxygen groups is low in most rubber-grade carbon blacks, and varies little from one grade to another, so the effect of the surface chemistry is generally minor.

Fumed silica, like carbon black, is produced in a flame (of silicon tetrachloride, hydrogen, and oxygen); thus its morphology resembles that of carbon black, though the aggregates may tend to be less tightly clustered than those of carbon black. Its surface contains hydroxyl (silanol) groups, in a concentration of 2–4.5 groups per nm^2; some nonpolar siloxane (Si–O–Si) groups are also present. The principal use in the elastomer industry is as a reinforcing agent for silicone rubber; the high cost of fumed silica precludes its usage in general purpose rubbers.

Precipitated silica is the second most widely used reinforcing filler, after carbon black. It is formed by controlled acidification of a solution of sodium silicate, stopping the reaction short of formation of massive silica gel. As with carbon black, precipitated silica exists in the form of aggregates, but these tend to be more highly clustered and some aggregates may resemble fragments of silica gel. The surface is covered with silanol groups, in a concentration of 8–12.5 per nm^2. Through these groups the aggregates tend to bond to each other, as well as to moisture; this makes the evaluation of some of the properties, such as bulkiness, difficult. Attempts to match the reinforcing performance of carbon black with precipitated silica have proven difficult owing to the different surface chemistry and morphology, but recently improved silicas, together with coupling agents (*vide infra*), have imparted excellent properties to tire trends ((Rauline 1991), and this is an active field of study at present.

Mineral fillers, especially clay and calcium carbonate, are widely used in the rubber industry. They function as low-cost extenders and also raise the modulus and stiffness, but are not generally considered to be reinforcing agents since the ultimate properties of mineral-filled rubber compounds are poor. Some improvement in reinforcement can be achieved by surface treatment so as to improve interaction (bonding) with the elastomer (Dannenberg 1982).

Short fibers are used to a limited extent, generally together with a reinforcing filler, to increase the modulus and dimensional stability, as well as to impart anisotropy of properties where this is desired. These fibers may be of glass, cellulose, carbon, or organic polymers, and are generally precoated with a binder to aid in dispersion and to improve bonding to the elastomer.

2. Mixing and Dispersion: State of Filler in Compound

Reinforcement requires good dispersion of the filler. Because of the high viscosity of elastomers, this requires equipment of high power, imparting a great deal of shearing energy, most of which is converted to heat. Cooling is relatively ineffective in large-scale equipment, so care must be taken to avoid high temperatures at which "scorch," or premature crosslinking, would take place. Internal mixers, such as the Banbury mixer, are generally used, but two-roll mills may also be employed (see *Polymer Melt Mixing: Agglomarate Dispersion**).

In the filler as received by the rubber manufacturer the individual colloidal entities (aggregates) are associated by van der Waals forces (for carbon black) or hydrogen bonding (for silica) in the form of agglomerates, which are formed into tiny spherical pellets (of carbon black) or may be pressed to remove air and ground and sieved to facilitate shipping and handling (of silica). The first stage in mixing is the incorporation, in which elastomer is squeezed into the void space in the filler pellets or agglomerates and the filler is incorporated into the polymer mass. The next, and overlapping, stage is dispersion, in which the elastomer-filled agglomerates are broken down into the individual aggregates. At the same time distribution is taking place, resulting in uniformity on a scale of tens or hundreds of micrometers. At the loadings generally employed for rubber reinforcement, the dispersed aggregates remain in some contact with each other even after the dispersion step is carried to good reinforcement, and this contact may be increased by reagglomeration after mixing has been terminated. Interaggregate contact results in formation of a filler network which is manifested by high dynamic modulus at low strains and (in the case of carbon black) by high electrical conductivity. Reagglomeration and network formation are especially pronounced in silica, resulting in a large increase in stiffness of the uncured compound on standing.

As may be inferred from the above discussion, there are two types or aspects of dispersion: one (macrodispersion) in which the initial compact agglomerates are broken down into aggregates, and the second (micro- or separation dispersion) which refers to the separation between the aggregates. Poor macrodispersion, in which rubber-filled agglomerates or pellet fragments greater than about 2 μm in size are present, leads to poor ultimate properties, since the undispersed regions act as weak spots or stress concentration centers at which failure can be initiated. Macrodispersion is what is ordinarily referred to as "dispersion"; it can be measured by microscopic examination of thin sections, or by roughness of torn or cut or stretched surfaces. Separation dispersion can be judged, though not measured in an absolute sense, by the electrical conductivity (of carbon black filled compounds), or by the Payne effect (*vide infra*). Increased separation dispersion results in decreased hysteresis and heat generation during dynamic mechanical deformation of the compound.

The high shear and elongational flow during mixing can result in some breakdown of the aggregates themselves. Breakdown of carbon black aggregates results in decreased bulkiness or "structure," and since the properties of the compound depend on the structure of the aggregates *in the compound*, it is desirable to estimate this in quality control of the carbon black. This is the purpose of the CDBP test (*vide supra*) but since the extent of aggregate breakdown depends on the polymer and mixing conditions, there is not generally a 1:1 correspondence between breakdown on crushing and breakdown on mixing. Furthermore the crushing procedure tends to bring the DBP absorption values of different carbon blacks closer together, thus the *un*crushed DBP absorption remains of value for quality control.

3. Mechanical Behavior of Filled Vulcanizates

Reinforcement is manifested in various aspects of the mechanical behavior of the reinforced vulcanized (i.e. cross-linked) rubber compounds (Wolff and Wang 1993; Medalia and Kraus 1994). These will be discussed in a sequence corresponding to increasing complexity of their presumed mechanisms. The effects of carbon black will be reviewed first, followed by a comparison with silica.

3.1 Elastic Modulus

In general, a filler will increase the elastic modulus of a rubber compound. A hydrodynamic basis for this effect, based on the Einstein viscosity equation, is given in the widely used Guth–Gold Equation:

$$G' = G_0'(1 + 2.5\phi + 14.1\phi^2) \qquad (1)$$

where G' and ϕ are the modulus and volume fraction of the filler, respectively, and the subscript zero indicates the absence of filler. Equation (1) gives good results for systems with spherical particles, including N990 carbon black (a poorly reinforcing carbon black with low surface area and low structure, composed largely of individual spherical particles). However, reinforcing grades of carbon black give elastic moduli considerably higher than the predicted values. This can be accounted for, in the limit approached at high double strain amplitude (DSA), by considering that rubber is occluded within the convoluted carbon black aggregates and that this rubber acts to some extent as part of the aggregate rather than as part of the rubber matrix. By calculating the amount of occluded rubber from the DBPA, and assuming that one-half of it acts as part of the aggregate, the following expression can be derived for the effective volume fraction of filler, V:

$$V = 0.5\phi(1 + [1 + 0.02139(\mathrm{DBPA})]/1.46) \qquad (2)$$

Substitution of V for ϕ in Eqn. (1) has given good results for a number of carbon blacks of different structure and surface area, at both low and normal loadings (Medalia 1973). For example, the predicted value of G'/G_0' from Eqn. (1) is 2.02 for all carbon blacks at a loading of 50 phr or $\phi = 0.1945$, as normally used for reinforcement (the expression phr refers to the parts by weight of filler per 100 parts of rubber); whereas the value calculated for N568 (a carbon black of DBPA = 142) using V from Eqn. (2) is 3.81, in agreement with the experimental value.

3.2 Augmentation of Elastic Modulus by Network Formation

In reinforced rubber compounds, unlike gum (unfilled) rubber, the elastic modulus is generally dependent on the strain amplitude. G' approaches a limiting value at very low amplitude, which is higher than the limiting value approached at high DSA. The extent of this augmentation ($\Delta G'$) depends on the loading and surface area of the carbon black, but not significantly on structure (Medalia 1978). At a normal loading of a reinforcing black the augmentation can be substantial; for example, with N220 (iodine number 126) at 50 phr in SBR-1500, G' is 44% higher at a DSA of 0.5% than at a DSA of 10%, while with N568 (iodine number 47) the augmentation is only 13% (based on measurements at 25 °C and 0.25 Hz). The augmentation is less at higher temperatures, greater at higher frequencies.

The augmentation of elastic modulus—often referred to as the Payne effect—is attributable to the network of carbon black which is broken down at high amplitudes. Mixing procedures which improve the separation dispersion diminish $\Delta G'$; for example, a two-stage mix in the above formulation with N220 gives only a 29% augmentation. Confirmation of the network as the source of augmentation comes from the electrical conductivity, which is higher for higher surface-area carbon blacks, and decreases with improved separation dispersion.

3.3 *Effect of Carbon Black on Hysteresis and Heat Generation*

Conversion of mechanical energy to heat, or hysteresis, is an important aspect of the dynamic behavior of rubber products (Medalia 1978); for example, in tires (see *Rubber Tires**), where energy loss (rolling resistance) as well as temperature rise (heat build-up) are significant factors affecting performance. While unfilled elastomers exhibit significant hysteresis, in reinforced elastomers the contribution of the filler generally outweighs that of the elastomer. The hysteresis, H, is the mechanical energy loss per unit volume per cycle; related parameters include the phase angle, δ, and the loss modulus, G'', where $G'' = G' \tan \delta$.

The contribution of the filler to hysteresis is attributable to breakdown and re-formation of filler–filler bonds and thus of the filler network. At very low-strain amplitude tan δ is no greater than that of the gum rubber, because there is little breakdown of the network. With increasing strain amplitude, tan δ increases, as interaggregate bonds are broken and then re-formed in new positions. Tan δ reaches a maximum at approximately the midpoint of the decrease in G', and thereafter decreases somewhat, though not nearly to that of the gum rubber. Note that in comparing different compounds, H is proportional to tan δ when the compounds are cycled at the same energy input, such as in a rebound experiment or (approximately) in tire treads; under equal strain conditions, H is proportional to G'', while under equal stress conditions, H is approximately proportional to $(\tan \delta)/G'$.

As may be expected, tan δ is a function of $\Delta G'$. Under comparable mixing conditions, tan δ has been found to be a linear function of a loading–interfacial area parameter, $\phi\psi$, where ψ is the interfacial area; thus $\phi\psi = \rho S \phi^2$, where ρ is the density of carbon black (1.85) and S is the specific surface area. A theoretical explanation for the relation between $\Delta G'$ and $\phi\psi$ has been derived by Kraus and subsequently refined (Ulmer 1996), on the basis of the kinetics of breakdown and re-formation of aggregate contacts, assuming that each such act dissipates energy.

In order to reduce rolling resistance, new elastomers have been developed in which the polymer chains are terminated by functional groups which can attach to the carbon black surface and thus impart better separation dispersion (Ulmer *et al.* 1998).

3.4 *Stress–Strain Behavior: Stress Softening*

A common test for mechanical properties is the uniaxial extension of a dumb-bell shaped specimen at a constant rate of strain to its breaking point. In comparison with the gum rubber, the initial slope of the stress–strain curve of a reinforced compound is much higher, owing to the hydrodynamic effect and to the filler network.

As the sample is stretched, the network is broken, but strain amplification contributes to a rapid rise in stress. Strain amplification refers to the increased strain on the rubber molecules linking rigid filler particles (aggregates). At very high strains, typically around 300% elongation ($\varepsilon = 4$), the rate of increase of stress starts to diminish, owing to several factors including polymer chain breakage, dewetting of the polymer from the filler–polymer interface, and slippage of the rubber molecules along the interface, until finally the specimen breaks. The stress at 300% elongation, based on the initial, cross-sectional area, is referred to as the 300% modulus (σ_{300}) and is widely used as a measure of stiffness while the stress at break, based similarly, is the tensile strength and is often taken as a practical measure of reinforcement. The 300% modulus depends on structure (compare N326, N330, and N347 in Table 1); the low modulus of the low-area black (N650) may be due to dewetting or molecular slippage, or to decreased shielding of the occluded rubber in the larger void spaces of this larger particle-size black.

If a filled rubber is stretched to a certain strain ε_1, then retracted to zero strain, then stretched again, the second stress–strain curve lies below the first one, but rejoins it at ε_1. Subsequent stress–strain cycles are progressively lower but to a much lesser degree. The mechanism of this stress softening or Mullins effect is not fully understood but may involve the same factors involved at higher extensions, or it may be due primarily to incomplete elastic recovery of the polymer, enhanced by strain amplification.

3.5 *Ultimate Properties*

The tensile strength depends on the surface area of the carbon black (compare N234, N347, and N650 in Table 1, the dependence being greater in SBR than in NR where self-reinforcing crystallites develop on stretching. (Note that in wear of tires the strain rates and temperatures are so high that crystallization cannot occur, thus reinforcement by carbon black is as important with NR as with noncrystallizing rubbers.) Tensile strength and elongation-at-break also depend on effective loading (i.e., volume fraction of

Table 1
Stress–strain properties of vulcanizates with various carbon blacks at 50 phr loading.

			Natural rubber			SBR-1500		
Carbon black	Surface area[a]	DBPA	300% Modulus	Tensile strength (MPa)	Elongation-at-break (%)	300% Modulus	Tensile strength (MPa)	Elongation-at-break (%)
N234	117	125	15.8	25.9	480	19.2	28.1	405
N330	80	102	14.4	25.2	520	17.7	26.9	435
N326	86	72	11.2	28.0	580	15.0	27.6	485
N347	86	124	16.3	25.2	480	20.1	26.9	420
N650	37	122	14.4	22.4	475	16.3	22.1	450

[a] By CTAB adsorption.

carbon black plus occluded rubber), going through a maximum at 50 phr of a low-structure black (N326), while higher structure blacks are to the right of the maximum of both properties at this loading.

Despite decades of research there is no generally accepted theory for the effect of carbon black on ultimate properties. Specimen breakage (tensile strength) occurs by propagation of an adventitious flaw or crack, generally at the edge of the specimen, and thus is related to fatigue crack growth, which can be measured separately but with much greater experimental difficulty (Young 1990). One school of thought focuses on dissipation of strain energy by the above-mentioned processes, including stress softening, thus leaving less energy available for the creation of new surface (Hamed 1994). However, studies of rubber tearing have shown that when tear deviation is prevented, strengthening by carbon black is relatively small. This and other evidence has led the author to conclude that reinforcement is due to formation of strands of carbon black aggregates on stretching, which act as barriers to tear propagation (Medalia and Kraus 1994).

Wear resistance is of vital importance in many applications, especially tire treads. The principal mechanisms involved in wear of rubber are abrasive wear and pattern abrasion. In abrasive wear, sharp asperities make scratches in the direction of abrasion and dig out the rubber while in pattern abrasion, adhesion of blunt asperities forms ridges normal to the direction of abrasion which are continually torn off. Abrasive wear resembles wear of plastics and is favored by high hardness of the rubber compound, while pattern abrasion, which is unique to rubber compounds, predominates with softer compounds (Veith 1992). Some widely used test machines favor one type of wear, some another type. Actual driving conditions may involve a mix of mechanisms, including mechanochemical degradation (see *Elastomers in Oxygen and Ozone, Degradation of*[#]); thus results from abrasion test machines are generally not considered reliable predictors of tire wear, which instead is evaluated by actual road tests, although because of the different mechanisms involved these do not always correlate well with each other. In a general way it is well established that a high-surface area (“reinforcing”) black is essential for good wear. The effects of structure and loading are complicated by the large effect of these variables on hardness, which is related to modulus, and thus to the predominant abrasion mechanism. Considerable effort has gone into development of carbon blacks with modifications in secondary characteristics such as aggregate morphology, aggregate size distribution, and surface activity (*vide infra*) in order to obtain some improvement in roadwear or in the balance of roadwear with other properties.

3.6 Comparison of Silica with Carbon Black

Precipitated silica, owing to its high coverage of silanol groups, interacts less strongly than carbon black with hydrocarbon rubbers, and forms stronger interaggregate bonds (network) and (Wolff and Wang 1992). This leads to higher hysteresis which does not set in until higher strain amplitudes. At high elongations silica compounds have low modulus, presumably due to slippage or detachment of rubber molecules. This slippage is beneficial in some applications, such as in imparting resistance to cutting and chipping of off-the-road tires.

The situation can be completely reversed by addition of a coupling agent which bonds the rubber to the silica. A widely used coupling agent is bis(3-[triethoxysilyl]-propyl)tetrasulfide, which reacts chemically via the triethoxysilyl groups with the silanol groups of the silica surface, and attaches to the elastomer molecules by way of the sulfide groups during vulcanization. Silica compounds with this coupling agent have *lower* hysteresis than carbon black compounds as well as high tensile strength and 300% modulus. With development of improved grades of silica, a good balance of tire tread properties can be attained (Rauline 1990), which may be sufficient to compensate for the drawbacks of these systems such as high cost of coupling agent, high Mooney viscosity of the uncured compound, more sensitive curing systems, and low electrical conductivity (requiring addition of a conductive carbon black in tire treads).

4. Rubber–Filler Interaction

When carbon black is mixed into rubber some of the rubber becomes bonded to the filler. The bonding has been variously attributed to physical adsorption, to chemisorption, to free radical reactions of broken elastomer chains with active sites on the carbon black surface or with fractured carbon black aggregates, or to ionic reactions (with butyl rubber). In a good solvent the unbonded rubber can be extracted, while the bonded or so-called bound rubber and filler form a coherent mass which swells considerably but does not dissolve (at room temperature). (However, most of the bound rubber can be extracted at elevated temperatures, strongly supporting the physical adsorption mechanism (Wolff and Wang 1992), with a range of adsorption energies.) The amount of bound rubber has been used as a measure of filler "activity"; however, only small differences are found in specific activity, or bound rubber per unit surface area. The amount of bound rubber increases with surface area and (to a small extent) with structure, thus for a given type of filler it correlates with the degree of reinforcement.

Precipitated silica gives considerably less bound rubber than a carbon black of similar surface area (e.g., 32% vs. 52%, with natural rubber at 50 phr loading). With graphitized carbon black, in which surface imperfections and (to a large extent) layer edges have been removed by heating to 2700 °C, the amount of bound rubber is very small (<5%). The mechanical behavior with a graphitized black is similar to that with silica: low 300% modulus, high elongation-to-break, and fairly high tensile strength (23.4 MPa vs. 26.2 MPa) for the ungraphitized black (Medalia and Kraus 1994). Formation of bound rubber is especially important with butyl rubber, where it is actually facilitated by oxygen groups on the carbon black surface.

The polymer adjacent to the carbon black surface is to some extent immobilized, as shown by NMR measurements. Severe restriction of motion is confined to a layer ~0.5 nm thick, while a further layer estimated at 3 nm is affected to a lesser extent; thus T_g (glass transition temperature) of the compound is raised by no more than 2–3 °C. The immobilized rubber must not be confused with *bound* rubber, since the long-chain molecules of the latter extend from the filler surface (where the bonded segments are immobilized) far into the polymer matrix, where they are unaffected by the filler surface.

5. Summary

Reinforcement is necessary for the applications, including tires, in which most rubber is used. Carbon black is the pre-eminent reinforcing agent, but precipitated silica is of increasing importance. Reinforcing fillers increase the elastic modulus by a hydrodynamic effect, in which the filler volume is increased by rubber occluded within the filler, and (at low strain amplitude) by a filler network. Breakdown and re-formation of this network on strain cycling is a hysteretic process which causes conversion of mechanical energy to heat. Improvement of ultimate properties (e.g., tensile strength, wear of tires) is an important aspect of reinforcement which is still poorly understood and thus represents a challenge for future research.

See also: Elastomers, Ceramic-modified*

Bibliography

Dannenberg E M 1982 Filler choices in the rubber industry. *Rubber Chem. Technol.* **55**, 860–80

Hamed G R 1994 Molecular aspects of the fatigue and fracture of rubber. *Rubber Chem. Technol.* **67**, 529–36

Medalia A I 1973 Elastic modulus of vulcanizates as related to carbon black structure. *Rubber Chem. Technol.* **46**, 877–96

Medalia A I 1978 Effect of carbon black on dynamic properties of rubber vulcanizates. *Rubber Chem. Technol.* **51**, 437–513

Medalia A I, Kraus G 1994 In: Mark J E, Erman B, Eirich F R (eds.) *Science and Technology of Rubber*, 2nd edn. Academic Press, New York

Rauline R 1991 *Patent application* EP 0501 227 (to Michelin)

Ulmer J D 1996 Strain dependence of dynamic mechanical properties of carbon black-filled rubber compounds. *Rubber Chem. Technol.* **69**, 15–47

Ulmer J D, Hergenrother W L, Lawson D F 1998 Hysteresis contributions in carbon black-filled rubbers containing conventional and tin end-modified polymers. *Rubber Chem. Technol.* **71**, 637–67

Veith A G 1992 A review of important factors affecting treadwear. *Rubber Chem. Technol.* **65**, 601–58

Wolff S, Wang M-J 1992 Filler–elastomer interactions. Part IV. The effect of the surface energies of fillers on elastomer reinforcement. *Rubber Chem. Technol.* **65**, 329–36

Wolff S, Wang M-J, Donnet J-B, Bansal R C, Wang M-J 1993 *Carbon Black*, 2nd edn. Marcel Dekker, New York

Young D G 1990 Application of fatigue methods based on fracture mechanics for tire compound development. *Rubber Chem. Technol.* **63**, 567–81

A. I. Medalia
Newton, Massachusetts, USA

Electronic Packaging: Conductive Adhesives

Electrically conductive adhesives perform two main functions. Like other adhesives, these materials provide a mechanical bond between two surfaces. In addition, they electrically interconnect the two bonded surfaces. This dual functionality is achieved by incorporating electrically conductive particles into a dielectric adhesive matrix. Conductive adhesives are used in electronic packaging applications when constraints such as process temperatures, stress levels, metallization options, contact pad pitch, or cost require an alternative to soldered interconnections.

1. Types of Conductive Adhesives

There are two principal types of conductive adhesives that are discussed in this article: isotropically conductive adhesives (ICAs), which conduct electricity uniformly in all directions, and anisotropically conductive adhesives (ACAs), which conduct electricity only in one direction (i.e., the direction perpendicular to the bond line). Conductive adhesives are the subject of many conferences (e.g., IEEE 1998) and publications. For a thorough review of conductive adhesives the reader is referred to Liu (1999).

1.1 ICAs

ICAs are prepared by the addition of conductive particles such that chains of particle-to-particle contacts span the material in all directions. For most common applications epoxies containing approximately 30 vol.% silver flakes are used which exhibit electrical conductivities of $1 \times 10^{-4}\,\Omega\,\mathrm{cm}$.

Increasing the concentration of metal particles in an insulating adhesive matrix changes the electrical properties of the composite in a discontinuous way. As the particle concentration begins to increase no significant bulk change occurs, because the particles are isolated from one another within the dielectric. Higher concentrations lead to the formation of agglomerates such that electrical conduction becomes statistically probable over a distance defined as the short-range percolation coherence length, ξ. At a critical concentration, called the percolation threshold (p_c), $\xi \to \infty$ and the composite becomes isotropically conductive.

The concentration of metal particles required to achieve p_c depends primarily upon the particle shape or aspect ratio (Bigg 1986). Increasing the aspect ratio from 1 (for spherical particles) to over 700 (for stainless steel fibers) decreases the percolation threshold from 38 vol.% to 1 vol.%. Other factors, such as the particle size distribution and material fabrication and processing techniques, can also have a significant impact (Lyons 1991). Most ICAs employ metal particles in flake form. The relatively high aspect ratio of flakes reduces p_c, which permits the formulation of adhesives with good rheological properties. Other techniques for exceeding p_c at low metal concentrations include the use of metal particles that are significantly smaller than discrete adhesive particles. In this way, the metal particles are segregated into the boundaries between adhesive domains.

The electrical conductivity of some ICAs develops only after the adhesives are thermally cured, even though the concentration of silver flake is above p_c. This is owing to the presence of a fatty acid coating applied to the silver surface during flake manufacture. Cure at elevated temperatures and/or the presence of reactive diluents removes the coating from the metal surfaces and renders the adhesive conductive (Opila and Sinclair 1985, Lovinger 1979).

1.2 ACAs

The conductivity of ACA materials is restricted to the Z-direction (perpendicular to the plane of the board) while electrical isolation is maintained in the X–Y plane. In general, ACA materials are prepared by dispersing electrically conductive particles in an adhesive matrix at a concentration far below the percolation threshold. The concentration of particles is controlled such that enough particles are present to assure multiple electrical contacts between the assembled parts in the Z-direction while too few particles are present to achieve conduction via percolation in the X–Y plane. Because of the anisotropy, ACAs may be deposited over the entire contact region, greatly facilitating material application and increasing overall bond strength. These attributes make ACAs especially useful for fine-pitch electrical interconnections.

The most commercially significant ACAs are based on the single-particle bridging concept. Interconnections with pitches below 50 μm have been demonstrated with these materials. Figure 1 illustrates the principle of interconnections based on this technique. An ACA is applied to the substrate, covering an array of contact pads. After deposition, a component is aligned to the substrate and placed into the adhesive. Pressure and heat are applied simultaneously such that excess adhesive is displaced and single conductive particles are trapped between the substrate and component contact pads forming multiple discrete electrical interconnections. Shrinkage of the adhesive during cure, contraction of the adhesive during cooling, and robust adhesion at all surfaces maintain the compressive load on the particle–contact metallization interface that is essential for reliable electrical contacts. Spherical particles (aspect ratio = 1) are used to lower the probability of conduction in the X–Y plane. The use of narrow particle size distributions

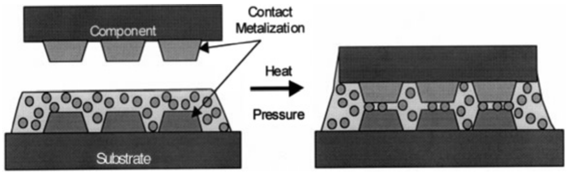


Figure 1
Schematic illustration of the ACA assembly process. Application of heat and pressure displaces excess adhesive to create redundant single-particle electrical interconnections in the Z-direction.

also reduces the probability of percolation in this direction at a given concentration (Lyons 1991).

Many adaptations of this basic approach have been developed to form ACA interconnections at fine (50 µm) pitches. Usually these approaches strive to minimize conduction in the X–Y plane. Detailed discussions of these, and other, processes are reviewed elsewhere (Liu 1999); a brief description of four techniques is presented here. In some methods, aligning metal particles in a magnetic or electric field enhances anisotropy and permits higher particle loadings. Other methods coat the conductive particles with a thin polymer dielectric layer. The coating fractures only when it is compressed between two contact surfaces, rendering the particle surface conductive. Some adhesive films have been developed with two layers of adhesive. One layer, containing the conductive particles, has a relatively high viscosity that constrains the ability of the particles to flow and form shorts. The second layer, which contains no particles, flows easily and fills voids. Another approach to fine-pitch interconnections deposits particles only on top of contact pads. This localization occurs through either selective deposition of an ACA, or photodefinition of a tacky adhesive layer on the pads, followed by adhering the particles to these tacky regions.

(*a*) *Conductive particle types for ACAs.* The type of particle incorporated into an ACA significantly impacts on the conductivity and reliability of the electrical interconnection. Three classes of particles, all of which are spherical in shape, are commonly used: rigid, compliant, and alloy particles, each of which is discussed separately. Contact theory (Holm 1967) provides a basis for understanding the electrical interconnection at the particle–contact metallization interface as individual particles can be considered to be asperities in Holm's model.

Rigid particles are fabricated from pure metals (e.g., silver), plated metals (e.g., silver-plated nickel), or metal-plated glass spheres. Particle selection depends on conductivity requirements and cost. During the assembly process sufficient pressure is applied so that particles penetrate into the contact surfaces forming intimate metal-to-metal contacts, as shown in Fig. 2(a). Plastic deformation occurs to a greater degree with softer contact metallizations.

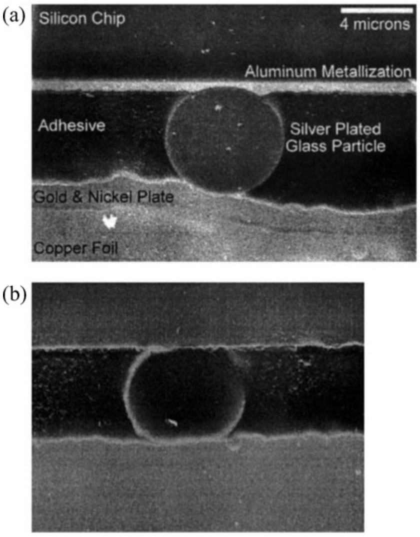


Figure 2
Scanning electron micrographs of ACA interconnections between silicon chips and flex circuits. (a) Rigid, silver-plated glass particle plastically deforms the relatively soft aluminum chip metallization. (b) Soft, nickel- and gold-plated polymer particle is plastically deformed between chip and flex circuit increasing the particle contact area.

Although they form stable interconnections, rigid particles can pose implementation problems. Some particulate materials with very broad particle size distributions contain numerous smaller particles that cannot form single-particle bridges. These "fines" may give rise to unintended conduction via percolation in the X–Y plane. Small concentrations of large, hard particles focus the force applied during assembly, resulting in excessive pressures that can crack brittle components.

Compliant particles are widely used as conductive fillers for ACAs and are typically composed of gold- and nickel-plated polymer spheres. During assembly, sufficient pressure is applied such that the particles themselves plastically deform and make contact over an area comparable to their initial diameter. An example of a deformed plated particle is shown in Fig. 2(b). As the size distribution of these particles is typically very narrow, with a range of sizes $<5\%$, most particles trapped between chip and substrate participate in the electrical interconnection to the same extent. The total force applied to an ACA joint, the number of contact pads (bumps), and the concentration of particles in the adhesive determine the amount of particle compression. Interconnect resistance and stability can be optimized by controlling the amount of compression (Shiozawa *et al.* 1995). The most stable values, for interconnections tested at 85 °C and 85% relative humidity, are reported for 40% particle compression levels. Higher compression levels can lead to cracking and reliability problems.

Factors affecting the resistance of interconnections formed with compliant particles are similar to those for rigid particles and include the number of particles, the contact area of each particle, the resistance of the particles, and the presence of surface oxides or other contaminants.

Solder particles form metallurgical bonds to both the substrate and component metallizations. This mechanism eliminates the need for the polymer adhesive to maintain the particles under a compressive load. The quality of the bond depends strongly on the interaction between the filler particles and the component and substrate metallurgies.

Savolainen and Kivilahti have reported much of the work on flip-chip interconnections with solder-filled ACAs (Kivilahti and Savolainen 1995, Savolainen and Kivilahti 1995). They found that the metallurgy of both the solder particles and the contact metals must be carefully chosen to avoid the formation of a brittle intermetallic layer within the joint. To circumvent this problem, they developed a transient melting technique using SnBi particles and tin or SnPb contact surfaces. Bismuth diffuses into the tin surfaces until the melting temperature of the progressively tin-rich particles increases and the particles solidify at the bonding temperature.

Contact resistance values are low owing to the large particle diameters that result from applying pressure to the soft solder particles. Oxides on the solder particle, chip, or substrate metallizations can prevent solder wetting and thus significantly increase resistance values. It is found, however, that no added flux is necessary to achieve good solder particle–surface wetting (Kivilahti and Savolainen 1995, Savolainen and Kivilahti 1995). Other studies have found the addition of flux to be advantageous (Pennisi *et al.* 1992).

2. Adhesive Matrices

The intimate particle–substrate contacts formed during assembly are maintained in the final interconnection by curing the adhesive. The adhesive serves several functions including: (i) bonding a component to the substrate through chemical and/or physical processes, (ii) protecting the exposed surfaces from environmental degradation, and (iii) maintaining continuous contact force at the particle–contact metallization interface. The higher coefficient of thermal expansion (CTE) of the adhesive matrix, compared to that of the particles, exerts a load on the contacts after cooling the assembly to ambient temperature (Lai and Liu 1996). Shrinkage of the adhesive during cure provides additional compressive forces on the particle contacts. Creep of the matrix would relax these loads. This explains, in part, why thermosetting matrices, which are less susceptible to creep owing to their high cross-link density, outperform thermoplastic-based conductive adhesives.

Two basic types of adhesives are used to formulate both ICAs and ACAs: thermoplastic and thermosetting. Thermoplastic adhesives are high molecular weight molecules that are rigid at temperatures below the glass transition temperature (T_g) of the polymer. Above this temperature, polymer flow can occur. When using this type of material, assembly temperatures must exceed T_g to achieve good adhesion. Thus, T_g must be sufficiently high to avoid polymer flow during use conditions, but must be low enough to prevent thermal damage to the electronic circuits during assembly. The principal advantage of thermoplastic adhesives is the relative ease with which the interconnection can be disassembled for repair operations. Block copolymers, such as styrene–butadiene–styrene have been used as the adhesive matrix. Newer systems employ linear polymers that can be cross-linked, improving performance.

Thermosetting adhesives are applied as mixtures of a curing agent with unreacted monomers of relatively low viscosity. Curing the mixture, typically with heat, causes a chemical reaction between curing agent and resin to form a three-dimensional cross-linked molecular network with a high T_g. Epoxies are the most widely used matrices for thermosetting adhesives owing to their ability to form robust adhesive bonds to a variety of surfaces over a wide range of environmental conditions. Details concerning the chemical formulation of conductive adhesives are reviewed elsewhere (Lyons and Dahringer 1994). Their tenacious properties can interfere with disassembly and interconnection repair steps.

Regardless of the type of adhesive matrix used, good adhesion is essential. This requires clean surfaces that can form chemical interactions with the adhesive. In addition, thorough wetting of the surfaces by the adhesive is required.

3. Assembly

The different conduction mechanisms of ICAs and ACAs mandate very different approaches to using these materials. ICAs require no external application of pressure during cure; however, the material must be placed only where needed. ACAs do require the simultaneous application of pressure during cure, but they may be applied over the entire surface to be bonded without creating unwanted short circuits.

Conductive adhesives come in two forms, films and pastes, regardless of the conduction mechanism involved. Films are formed from polymers or resins that are solid at room temperature. Using liquid polymers, solvents, and/or elevated temperatures, a mixture of the adhesive system with conductive particles is spread onto a sheet, then dried, partially reacted, or cooled to form a freestanding film. Pastes are typically manufactured with low-viscosity thermosetting adhesive systems. Solvents are avoided in all adhesive systems as voids form during volatilization.

Assembly issues for the two classes of conductive adhesives are discussed separately.

3.1 ICA Assembly

The first step in assembling components with ICAs is dispensing the adhesive in the correct location. Commonly used techniques include syringe dispensing, stencil and screen printing, and dipping. Film adhesives are used for larger bond areas and employ manual or mechanical placement tools.

Syringe dispensing is the most commonly used method for die attach adhesives. An array of needles creates a specific pattern of adhesive on the substrate (lead frame). The adhesive pattern is designed to minimize the entrapment of air when the die is placed into the adhesive. Automated equipment controls the dispense time and pressure resulting in reproducible adhesive volumes.

Stencil and screen printing processes are best suited for assembly processes occurring over large areas. For example, surface mount assembly of components on flex or rigid PWBs is best accomplished with stencil printing an ICA as large numbers of adhesive dots of various sizes can be easily and reproducibly accommodated. Screen printing of circuit patterns on flexible PWBs with ICAs formulated as inks (low-viscosity pastes) is another commonly used process.

Interconnection of fine-pitch components with ICAs imposes severe tolerances on all application processes. The volume of adhesive necessary to interconnect very fine-pitch surface mount components is difficult to optimize as the volume required to ensure electrical isolation between adjacent leads may be insufficient to ensure the mechanical integrity of the bond, especially during drop tests (Zwolinski *et al.* 1996).

The use of ICAs for flip-chip interconnections is beginning to gain acceptance. A process developed by Matsushita (Kumano *et al.* 1999) employs ICA-coated stud bumps. Bumps formed with a wire bonder on integrated contact pads ($\lambda = 75\,\mu\text{m} \times 75\,\mu\text{m}$) are dipped into an ICA paste to a controlled depth. The paste coats the stud bumps and provides sufficient adhesive to comply with any nonplanarity resulting from either inconsistent bump heights or substrate flatness variations. Since the coating volume is low, insufficient adhesive is present to cause any shorting. To create a robust mechanical bond, an underfill adhesive is required to fill the gap between chip and substrate and adhere to the exposed surfaces.

3.2 ACAs

ACAs are applied using the same methods as those described for ICAs; however, tolerances on the application area are significantly relaxed as the ACA can be applied to the entire interconnection area. This has numerous advantages, especially for fine-pitch interconnections, as sufficient adhesive is present to ensure robust adhesive bonds. Film-based ACAs (ACAFs) require special machines to cut, align, and tack the adhesive onto the substrate. Film thickness must be carefully selected to meet the requirements imposed on the assembly by the height of the contact pads, as well as the conductive particle diameter in the adhesive. Insufficient adhesive results in voids at the bondline whereas excessive adhesive can be difficult to displace thereby preventing particles from contacting the metal surfaces.

All ACA assembly processes require the application of pressure during adhesive cure. Maintaining alignment between component and substrate, while simultaneously applying heat and pressure, presents a challenge to developing a cost-effective assembly process.

For flip-chip bonding the alignment machine is commonly used to perform several functions including alignment of the chip to the substrate, application of sufficient energy to cure the adhesive, and application of pressure uniformly across the chip. Because of the multiple requirements and tight tolerances, machines are expensive to own and maintain. Cure time reduces throughput and significantly increases machine operating costs. This technique is widely used during the manufacture of LCD displays for either interconnecting bare die directly to the glass substrate or for bonding TAB packaged ICs to the glass substrate. The costs of assembling flip-chip interconnections with ACAs have limited the implementation of ACA materials to products with special materials or other constraints. Lower cost and higher throughput assembly processes are required to expand the use of ACA materials to applications beyond interconnections to glass substrates.

A new approach to using ACAs has been described (Lyons *et al.* 1996a, 1996b). In this process, the cure step is shifted from the alignment machine to a batch-curing fixture (BCF) where a conformal silicone bladder applies pressure uniformly to all assemblies. Many types of components, ranging from bare die to SMT packages, can be assembled with ACAs using this approach. A cross-sectional view of the BCF apparatus is shown in Fig. 3. Machines for evaluating the BCF process are commercially available (Fancort Industries, West Caldwell, NJ).

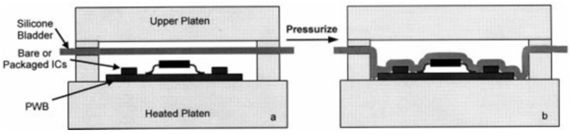


Figure 3
Schematic cross-section of the batch-curing fixture: (a) after clamping the fixture to create a sealed sample compartment; (b) after pressurizing the upper chamber, forcing the bladder into conformal contact with the bare and packaged ICs.

When implementing the BCF process, the first step is to apply the ACA to the substrate. A paste adhesive maximizes the throughput of the BCF process for two reasons. Paste adhesives can be formulated with specific rheological properties such that components placed into an ACA paste are held securely owing to their "tacky" nature. This eliminates the need to supply heat to the component in the alignment machine. The second reason is that stencil printing paste is an inherently less expensive process than cutting, aligning, and tacking film-based adhesives. Film materials can be used but component tacking in the alignment machine is required.

After printing, components are aligned and placed into the paste. The populated circuits are transported from the placement machine to the BCF for cure. The BCF sample chamber can be designed to accommodate single or multiple circuits, depending on sample size and process flow.

In the BCF, the populated circuits are placed on a heated plate, as shown in Fig. 3(a), and a stretched silicone rubber bladder is placed above the circuits to create a sample chamber. Pressure is initially applied to the interconnections by first evacuating the sample chamber (Fig. 3(b)). Ambient pressure above the bladder forces it into conformal contact with the populated circuits. Compressed gas (100–500 psi (6.9×10^5–3.4×10^6 Pa)) is introduced above the bladder during the cure cycle to ensure intimate particle–metal contact. Pressure can be maintained during cooling of the samples and is vented before opening the BCF and removing the completed circuits.

The chip-to-flex alignment achieved during the initial placement step is maintained to within $\pm 5\,\mu\text{m}$ after final cure. This ability to reproduce closely initial alignment accuracy is achieved, despite transporting the populated circuits from the alignment machine to the BCF, owing to the tacky nature of the ACA paste as well as the compliance of the silicone bladder. The bladder ensures that pressure is uniformly transmitted to the silicon chips regardless of substrate–bladder coplanarity.

The principal advantage of the BCF process is that it frees the alignment machine to *only* place die. Thus, throughput in the machine is significantly improved, approaching the rates achievable for flip-chip solder interconnections (only the increased alignment accuracy slows the machine). Equipment costs are reduced, as the machine need not be constructed with the ability to apply significant pressure levels, and chip–substrate coplanarity requirements are relaxed.

Assembly considerations for SMT components are similar to those described for flip-chip processes. A thorough, step-by-step process description for very fine-pitch flex to PWB substrates has been published (Bruner 1995). Throughput is an even greater concern, as placement times for SMT components are even shorter than those for bare die. A BCF process for SMT components has been developed to address these concerns.

4. Reliability of ACA Interconnections

Optimization of materials, processes, and surfaces results in robust interconnections for both ICAs and ACAs. The reliability of these conductive adhesive interconnections depends on the specific formulation, process, substrate, and component combination. No *a priori* assumptions can be made, as each system must be evaluated separately. The reliability of conductive adhesive interconnections has been thoroughly documented for significant applications including die attach (ICAs) and LCD assemblies (ACAs). As experience with these relatively new materials broadens, and as fine-pitch requirements increase, the use of conductive adhesives is expected to grow.

Bibliography

Bigg D M 1986 Electrical properties of metal-filled composites . Bhattacharya S K (ed.) *Metal Filled Polymers*. Dekker, New York, p. 165

Bruner D M 1995 Z-axis adhesive film interconnect using 150 micron and 75 micron pitch flex circuits. *Int. J. Microcircuits Electron. Packag*. **18** (3), 311–8

Holm R 1967 *Electrical Contacts; Theory and Application*, 4th edn. Springer, New York

IEEE 1998 *Proc. 3rd Int. Conf. Adhesive Joining and Coating Technology in Electronics Manufacturing*. IEEE, Piscataway, NJ

Kivilahti J, Savolainen P 1995 Anisotropic adhesives for flip-chip bonding. *J. Electron. Manufact*. **5** (4), 245–52

Kumano Y, Tomura Y, Itagaki M, Bessho Y 1999 Investigation of chip-on-flex application using SBB flip-chip technique. *Proc. 1999 Int. Symp. Microelectronics*. Chicago, pp. 137–42

Lai Z, Liu J 1996 Anisotropically conductive adhesive flip-chip bonding on rigid and flexible printed circuit substrates. *IEEE Trans. CPMT, Part B* **19** (3), 644–60

Liu J (ed.) 1999 *Conductive Adhesives for Electronics Packaging*. Electrochemical Publications, UK

Lovinger A 1979 Development of electrical conduction in silver-filled expoxy adhesives. *J. Adhesion* **10**, 1–15

Lyons A M 1991 Electrically conductive adhesives: effect of particle composition and size distribution. *Polym. Eng. Sci.* **31**, 445–50

Lyons A M, Dahringer D W 1994 Electrically conductive adhesives. In: Pizzi A, Mittal K L (eds.) *Handbook of Adhesive Technology*. Dekker, New York, pp. 565–84; and references therein

Lyons A M, Hall E, Wong Y-H, Adams G 1996a A new approach to using anisotropically conductive adhesives for flip-chip assembly. *IEEE Trans. CPMT, Part A* **19** (1), 5–11

Lyons A M, Kammlott G W, Wong Y-H, Adams G 1996b A batch curing process for the assembly of flip chip interconnections with anisotropically conductive adhesive pastes. *Proc. Adhesives in Electronics '96 Conf.*, Stockholm, pp. 68–75

Opila R L, Sinclair J D 1985 *Ann. Proc. Reliab. Phys. Symp.* **23**, 164

Pennisi R W, Papageorge M V, Urbish G F 1992 *US Pat.* 5 136 365

Savolainen P, Kivilahti J 1995 Electrical properties of solder filled anisotropically conductive adhesives. *J. Electron. Manufact.* **5** (1), 19–26

Shiozawa N, Isaka K, Ohta T 1995 Electric properties of connections by anisotropic conductive film. *J. Electronics Manufact.* **5** (1), 33–7

Zwolinski M, Hickman J, Rubin H, Zaks Y, McCarthy S, Hanlon T, Arrowsmith P, Chaudhuri A, Hermansen R, Lau S, Napp D 1996 Electrically conductive adhesives for surface mount solder replacement. *Proc. Tech. Program, SurfaceMount International*

A. M. Lyons
Lucent Technologies, Murray Hill, New Jersey, USA

Electronic Packaging: Elastomer Conductive Polymers

Conductive elastomers (CEs) have been used for many years in switches in automobile horns and door open sensors, and automobile ignition wires are made of this material. CEs have found wide use in applications involving dissipation of static electricity. This includes many applications in hospital operating rooms where a spark could ignite highly explosive anesthetics. Automobile and truck tires, especially for vehicles carrying explosives, are made conductive to eliminate sparking. Many integrated circuit (IC) chips can be destroyed by static discharge. To prevent this, CE devices are used to electrically connect the chip leads during shipping and also ground those handling the chips. CEs have been used extensively in electro-magnetic shielding gaskets in military and nonmilitary applications.

CEs are used because of their ability to conform to substrate contours, their ease of fabrication into a wide variety of shapes and forms, and their versatility in electrical and mechanical properties.

1. Composition and Theory

A CE is generally produced by the incorporation of conducting particles in a rubber matrix. Inherently conductive polymers, such as polyacetylene, show promise for future applications, but their lack of stability and processing problems have limited their usefulness.

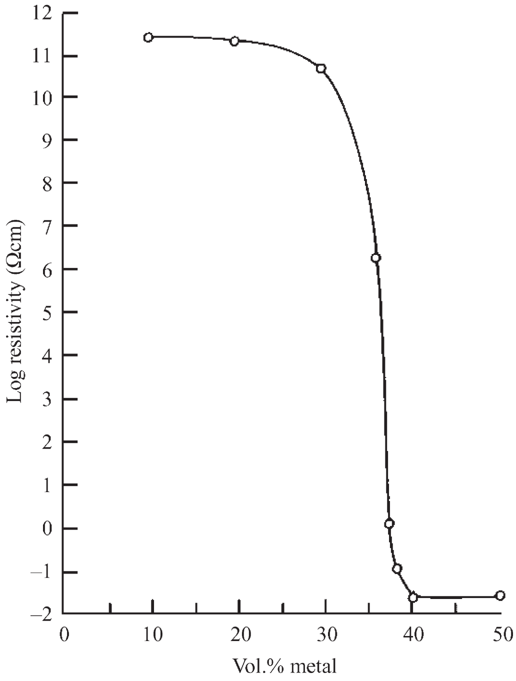


Figure 1
Typical electrical resistivity (plotted as the logarithm) of a metal particle-filled plastic as a function of percentage filler.

The conductivity of conductive particle-filled silicone rubbers results from intimate contact and electronic tunneling between particles. As shown in Fig. 1, there is a loading at which the conductivity will increase (resistivity will decrease) very rapidly with a small increase in filler content. This is the result of a process called percolation, in which a continuous chain of particles is required to produce conductivity (Wessling 1986). The conductivity of the filled rubbers varies with the metals used, shape and size of the particles, fabrication techniques, and cure time and temperature. The properties of the elastomer matrix also have a profound influence on the immediate and long-term properties of the material.

1.1 Matrix Composition

The elastomer matrix is generally silicone rubber because of its electrical and physical stability (Buchoff 1973). It maintains its properties over a wide range of temperature, humidity, radiation, and chemical environments. The modulus of a typical dimethyl silicone rubber varies little from $-45\,^{\circ}C$ to $150\,^{\circ}C$. It can

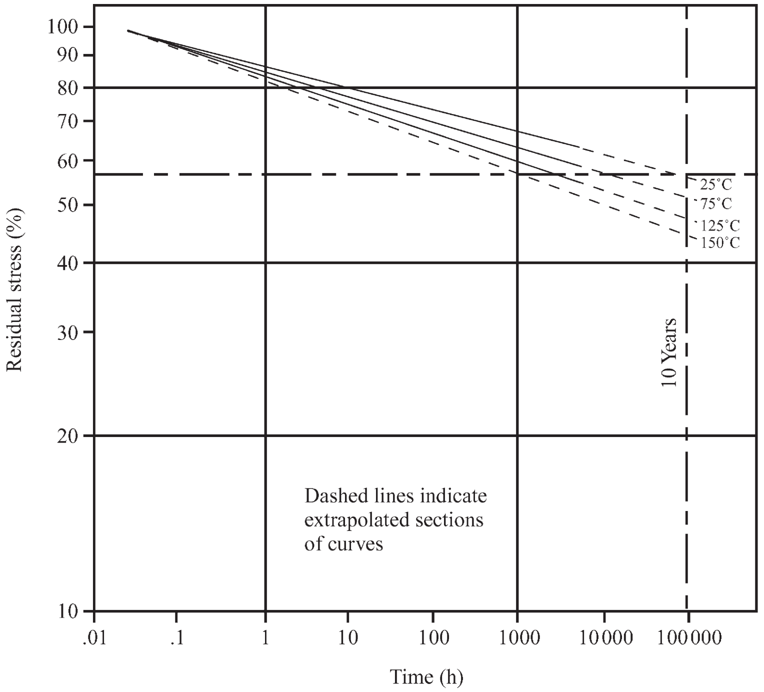


Figure 2
Logarithm of the percentage residual stress of a silicone elastomer held at constant strain as a function of logarithm of time at different temperatures.

withstand a temperature of 175 °C continuously and 300 °C for a few hours. Phenyl methyl silicones have the same high-temperature properties but remain flexible below −75 °C. Even when saturated with water, silicone rubbers maintain a volume resistivity greater than $10^{12}\,\Omega\,\text{cm}$. They have dielectric constants around 2.8 and very low loss tangents. Hydrocarbons and halogen-containing hydrocarbons cause silicone rubbers to swell, but will not chemically change them. When the solvent is removed from the silicone, the properties of the silicone rubber are the same as they were before solvent exposure.

Particle-filled silicone rubber exhibits classic stress relaxation (Buchoff 1983). That is, when the filled rubber is deflected and held at a constant strain, the stress decays with time as shown in Fig. 2. There is a linear relationship between logarithm of remaining stress and logarithm of time. This allows projection of the remaining contact pressure over a wide range of temperature and time. Extrapolating, this curve shows that, when held at a constant strain, silicone rubber loses only 50% of the original stress in 50 years at 25 °C (77 °F). There is also a linear relationship between logarithm of resistance and logarithm of time as shown in Fig. 3. The rate of this decay—termed relaxation—increases as the temperature increases. When the material is deflected, its volume increases slightly, separating the particles. Also the particles tend to reorient, perpendicularly to the imposed force. These changes result in a decrease in volume conductivity. Over time, as the elastomer relaxes, the particles move closer together and they gravitate toward their original orientation. Thus, the conductivity increases. Because of the initial higher resistance, it is desirable to wait 15–20 seconds after deflection before imposing a voltage across the material.

1.2 Conductive Fillers

The most widely used conductive fillers are carbon, silver, nickel, stainless steel, and silver-plated copper, graphite, nickel, aluminum, and glass beads. Powder, flakes, beads, and filaments of various sizes and shapes are used (Roos-Kozel and Casavecchia 1984).

The initial conductivity of elastomers containing silver-coated copper is almost as high as those containing pure silver, but at a fraction of the cost of silver. Unfortunately, the copper migrates through the silver layer, where it oxidizes and sharply increases the resistivity. Elevated temperature greatly increases the rate of migration. Silver-coated aluminum particles are more stable than silver-coated copper particles, but they produce a lower conductivity.

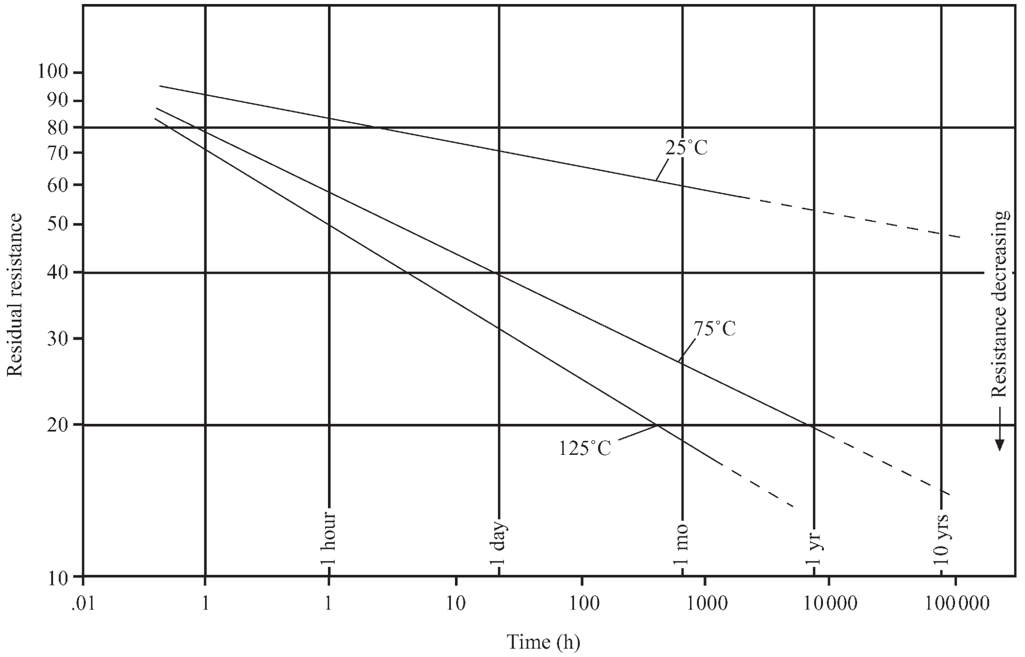


Figure 3
Logarithm of percentage residual electrical resistance of a metal particle-filled silicone elastomer held at constant strain as a function of logarithm of time at different temperatures.

Carbon filler is employed if a volume resistivity greater than 1 Ω cm is required. Silver and silver-plated particles yield volume resistivities in the range 10^{-4}–10^{-1} Ω cm.

2. *Uses of CEs in Electronic Applications*

CEs are used in electronic equipment such as shielding gaskets, pressure transducers, switches, keyboards, test contacts, shorting pads, and thermal coupling. A large and rapidly growing use is in connectors in digital watches, calculators, displays, chip carrier test contacts, military radar, cellular telephones, pagers, medical devices (Galton 1975, Pelt 1989), and space electronic gear (Markstein 1992).

3. *Elastomeric Connectors*

Elastomeric connectors are zero insertion force devices, i.e., there is no force used to install the connectors but pressure is necessary to establish and maintain electrical contact. A holder usually positions the element(s) and aligns the substrate contact pads (see Fig. 4). Pressure is applied with screws, rivets, headed plastic stubs, C-sections, etc. Often, the required deflection is produced during the assembly of a product (Buchoff 1987).

3.1 Advantages of Elastomeric Connectors

The continued pressure at the elastomer–pad interface effectively seals out moisture and other atmospheric contaminants. Even in highly corrosive environments the elastomer seal prevents the formation of insulating films that would increase resistance. This has been shown in a test in which mated elastomeric connectors are exposed to an atmosphere containing 5 ppm each of hydrogen sulfide, sulfur dioxide, and nitric oxide at 35 °C (95 °F) and 75% relative humidity for a 1000 h. There is no increase in resistance and the copper contact pads touching the connector elements are as shiny as before the test (Buchoff 1974).

Use of elastomeric connectors eliminates the heat of soldering and the necessity of removing flux and other contamination. Also, disassembly and reassembly is greatly simplified, and repair and upgrading of electronic equipment is fast and without damage to other components. Delicate parts, such as liquid crystal displays (LCDs), are shielded from shock and vibration. Also, the range of usable deflection of the elastomeric elements reduces the assembly tolerance requirements.

3.2 Types of Elastomeric Connectors

There are several configurations of elastomeric connectors employing conductive particles (Carlisle 1983).

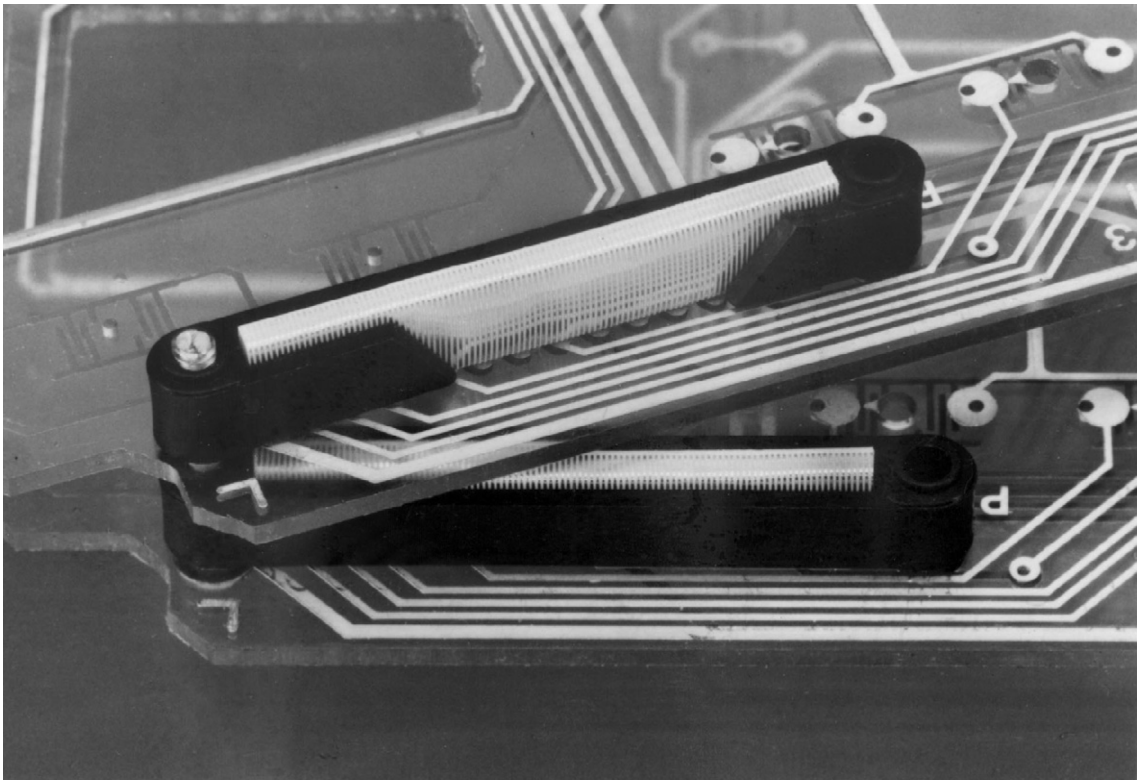


Figure 4
Elastomeric metal ribbon connector element in a holder.

These include: (i) layered elements, (ii) buttons of conductive rubber molded into holes in plastic supports, (iii) conductive rubber inclusions into a nonconductive elastomeric support, and (iv) sheets of rubber containing metal particles at a concentration too low to produce isotropic conductivity.

(*a*) *Layered elements*. Layered elastomeric elements (LEEs) consist of alternating segments of conductive and nonconductive rubber, as shown in Fig. 5. LEEs are redundant components, i.e., the layers are much narrower than the substrate pads. Several of the conductive layers make contact with each pad regardless of the element position. This eliminates the need to align the conductive elements with the contacts, simplifying installation. The width of the conductive layers is typically 0.062 mm (0.0025 in) and they are on a pitch of 0.125 mm (0.005 in). Reliable contact is made to pads on 0.4 mm (0.016 in) or greater centers. Carbon-filled LEEs with a 0.040 mm (0.0016 in) pitch have been produced.

Carbon LEEs form the contact between the circuit boards and LCDs in the majority of digital watches (Buchoff 1974). The 100–5000 Ω resistance of the carbon-filled LEE is no problem because of the very high internal impedance of the LCDs. Where lower resistance is required the filler is silver. Resistances from 0.01 Ω to 0.5 Ω are common with silver-containing LEEs.

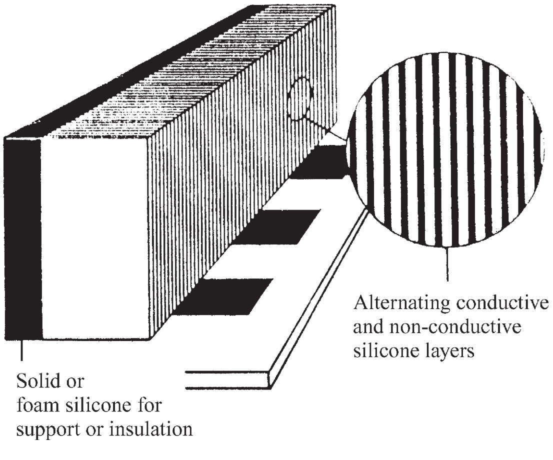


Figure 5
Layered elastomeric connector element.

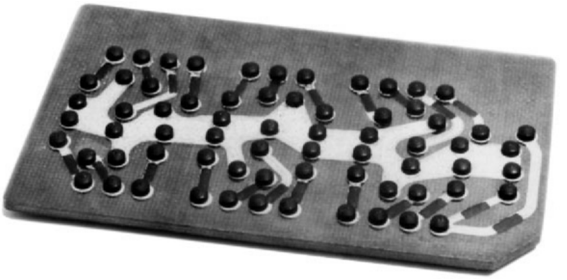


Figure 6
Interposer containing an array of conductive elastomeric button contacts.

(*b*) *Conductive buttons.* Conductive elastomer buttons molded into holes in carrier frames made the contacts in the earliest LCD watches. The frames containing the appropriate holes were molded in glass-filled nylon or glass-filled polyester. Thin Kapton sheets with punched holes also served as carriers. This configuration has also been used to make connections to pads for area matrices, as shown in Fig. 6. Pads having pitches as fine as 0.5 mm (0.020 in) have been connected. Conductive rubber buttons have been molded onto circuit paths, printed wiring boards (PWBs), flat cable, and flexible circuits.

Molded wiring boards containing conductive rubber contact balls have been produced where the hold-downs are molded as part of the board. Locating means for components are also incorporated, making automated assembly possible. Components can be replaced rapidly without disturbing the package.

A device has been developed to connect high I/O microprocessors to PWBs. Columns formed from a flexible conductive polymer material with embedded metallized particles make the contacts (Anonymous 1998).

(*c*) *CEs in nonconductive elastomer bodies.* An area of rapid growth is parts with conductive elastomers molded into nonconductive elastomer bodies (Maurer 1998). In one application, both silver and carbon LEEs are mounted onto a silicone rubber foam piece. The carbon LEE makes connections to a LCD and the silver LEE contacts the LCD heater. In another highly successful application, contacts are made between a PWB and a microphone. A nonconductive elastomer piece containing a CE insert forms the contact. The assembly is made by placing the elastomer part and the microphone into a positioning cavity followed by clamping.

(*d*) *Z-axis elastomeric sheets.* One of the earliest elastomeric connector elements used a silicone rubber sheet containing metal particles below the critical volume (Jin *et al.* 1988). That is, the particles are in too low a concentration to produce isotropic conductivity. When this sheet is sandwiched between two sets of pads, connections are established in the Z-axis between the opposing pads without creating conductive paths to adjacent pads. Z-axis material has been used in switches in automobiles for many years.

When the normally insulating Z-axis material is pressed, a conductive path is produced between two opposing pads. This type of switch has been used widely to sense door closure or activate the horn. It is also used in cameras to trigger camera shutters. The pads can be in a single row or over an area. One of the limitations of this approach is the dependence of the thickness of the sheet on the minimum separation of the pads. Generally electrical isolation between adjacent pads is maintained if the sheet thickness is less than 70% of the pad separation. This requirement leads to thin sheets, necessitating flat mating surfaces and high pressure to ensure low-resistance connections. The performance of this material can be improved by using silver-coated nickel particles that are vertically aligned in a magnetic field during production of the pad.

3.3 Elements Containing Metal Ribbons

In another family of elastomeric connector elements, thin, continuous metal strands run through, on one surface, or around the silicone rubber base (Buchoff 1983). The strands are wires or ribbons. Wires are usually stainless steel or nickel, often gold plated. The ribbons most often used are nickel, gold, or gold on nickel on copper.

Unlike LEEs, the metal traces of ribbon elastomeric elements (REEs) can be positioned accurately. Using REEs with ribbons on one flat surface of the silicone rubber core, contact has been made to pads on 0.1 mm (0.004 in) centers. These flat REEs produce minimum signal distortion at frequencies above 10 GHz when connecting sets of pads in the same plane separated by 0.125 mm (0.005 in). Position, width, and shape of the metal traces can be varied to contact pads on different pitches and/or produce specialized electrical properties. This configuration is used to connect circuit boards to bumped IC chips, circular IC chip probe boards, and high-frequency hybrid circuits.

Connections to medical implants with flat REEs operate well during long-term immersion in saline solution and have been implanted in several patients. This *in vivo* junction will simplify the implementation of advanced medical systems.

In another REE configuration, the metal traces run through the elastomeric body, making contact at the trace ends. The traces form a two-dimensional array as shown in Fig. 7. As there is a multiplicity of conductive paths, higher conductivity contacts are

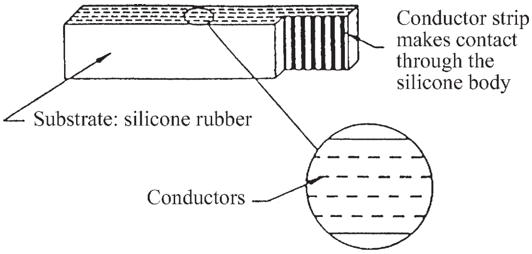

Conductor strip makes contact through the silicone body
Substrate: silicone rubber
Conductors

Figure 7
Elastomeric connector element with metal traces running through the silicone rubber body.

produced. Also, multiple rows of contacts can be made with a single elastomeric element. This type of connector has proved to be highly reliable in a phased array radar system.

In wrap-around REEs, the metal traces encircle the rubber core (see Fig. 4). Contact is established between opposing sets of contact pads when they are pressed against the traces. Generally, as with LEEs, the wrap-around REE makes redundant contact. The traces have been produced on centers ranging from 0.1 mm to 0.5 mm (0.004–0.02 in). Resistance ranges from 0.005 Ω to 0.05 Ω depending on the dimensions of the contact pads and composition of the traces.

By changing the material of the core and its shape, clamping pressure can be adjusted from 10 psi to 100 psi (69×10^3–690×10^3 Pa).

4. Future Developments

As electronic packages and components get smaller and the pin counts increase, the use of elastomeric connectors will increase. Portable devices such as cellular telephones, pagers, calculators, computers, and data acquisition terminals will benefit the most. Elastomeric connectors will allow compact military devices to be smaller, have more closely spaced contacts, and be more reliable (Buchoff 1989). Connections will be made directly to chips and chip-scale packages.

Conductive elastomeric elements will simplify, improve, and reduce the cost of test and burn-in fixtures (Buchoff 1980). The majority of test fixtures employ spring-loaded metal pins or cantilevered metal arms. At close contact spacing, the elements in both techniques are fragile and easily damaged and become misaligned. Contacts spaced on 0.1 mm (0.004 in) centers have been successfully probed with metal-on-elastomer elements. When it becomes necessary to replace metal contact elements, the entire socket must be replaced. In contrast to this, individual elastomeric elements are easily replaced without removing the sockets.

New applications will be found such as sensors, switches, and even active devices. CE shapes will be molded directly onto product cases and PWBs to form gaskets, conductive paths, pressure sensors, sockets, resistors, heaters, current limiters, and other devices. Conductive and nonconductive elastomers will be formed together to make complete assemblies.

See also: Electronic Packing: Conductive Adhesives*

Bibliography

Anonymous 1998 Thomas and Betts unveils new MPI technology, *SMT* Nov., 39–40

Buchoff L S 1973 Use of conductive elastomers in electronic connectors. *Proc. Ann. Connector Symp.*

Buchoff L S 1974 Conductive elastomers make small flexible contacts. *Electronics* (Sept.), 16

Buchoff L S 1980 Elastomeric connections for test and burn-in *Microelectron. Manufact. Testing* (Oct.)

Buchoff L S 1983 Surface mounting of components with elastomeric connectors. *Electronics* (June)

Buchoff L S 1987 Guidelines for designing elastomeric connectors into the system. *Connection Technol.* (Aug.)

Buchoff L S 1989 Elastomeric connectors for land grid array packages. *Connection Technol.* (April),

Carlisle B 1983 Solving problems with elastomeric connectors. *Machine Design* (Nov.), 24

Galton L 1975 Dramatic successes with electromedicine. *Parade* (Dec.), 14

Jin S, Mottine J J, Seger S G, Sherwood R, Tiefel T H 1988 Anisotropically conductive composite medium. *US Pat.* 4 737 112

Markstein H W 1992 Applications widen for elastomeric connectors. *Electronic Packaging and Production* (Nov.), 24

Maurer D J 1998 Elastomeric connectors in a new role. *Sensors* (Feb.)

Pelt D V 1989 A bandage with the power to heal. *Insight* (Feb.), 13

Roos-Kozel B L, Casavecchia 1984 Parameters affecting silver flake and powder performance in silver filled polymer. *Solid State Tech.* (Aug.), 167–71

Wessling B 1986 Electrically conductive polymers. *Kunststoffe* **76** (10), 930–6

L. S. Buchoff
TYCO, Fort Washington, Pennsylvania, USA

Electronic Packaging: Heat Sink Materials

Heat sinks play two key roles in electronic packaging: thermal management and mechanical support. They are used in packages and with printed circuit boards (PCBs). Traditional heat sinks were primarily aluminum, copper, or unalloyed blends of two metals, such as copper/tungsten or copper/molybdenum. Because of the shortcomings of these traditional heat sinks, there has been considerable work on new materials, primarily composites reinforced with fibers and particles. Advantages of these new materials include high thermal conductivities; low, controllable coefficients of thermal expansion; weight reductions; high strength and stiffness; and availability of net shape fabrication processes. In some cases the new materials are cheaper than the ones they replace.

1. Requirements for Heat Sink Materials

Heat sinks support electronic devices and provide a path for heat dissipation. For larger heat sinks, such as those used with PCBs, strength and stiffness are important properties (PCB heat sinks are also called thermal planes and cold plates). The continuing increase in packaging density has resulted in a need for materials with high thermal conductivities. In addition, to minimize thermal stresses that can cause component or solder failure, it is desirable that packaging materials have coefficients of thermal expansion (CTEs) matching those of the ceramic components, such as substrates and resistors, and semiconductors they support. Further, low density is desirable in many applications, especially transportable systems such as laptops, hand-held telephones, and avionics. Reducing weight also minimizes potentially damaging stresses resulting from shock loads that can occur during shipping and from other causes. Of course, low cost is also a key consideration. Traditional materials used in electronic packaging do not meet all of these requirements. In response to this need, new materials, especially composites, have been, and are continuing to be developed.

A composite material can be defined as two or more materials bonded together (Kelly 1988). Composites are nothing new in electronic packaging. For example, polymer matrix composites (PMCs) in the form of E-glass fiber-reinforced polymer PCBs are well-established packaging materials. Similarly, a variety of particles are added to polymers to reduce CTE and increase thermal conductivity, electrical conductivity, or both. These materials are particle-reinforced PMCs. Composites have been developed that provide unique combinations of properties making them outstanding candidates for heat sinks.

The four key classes of composite materials are PMCs, metal matrix composites (MMCs), ceramic matrix composites (CMCs), and carbon matrix composites (CAMCs), which include carbon/carbon composites (CCCs). PMCs, MMCs, and CCCs are the key composites of interest for thermal management.

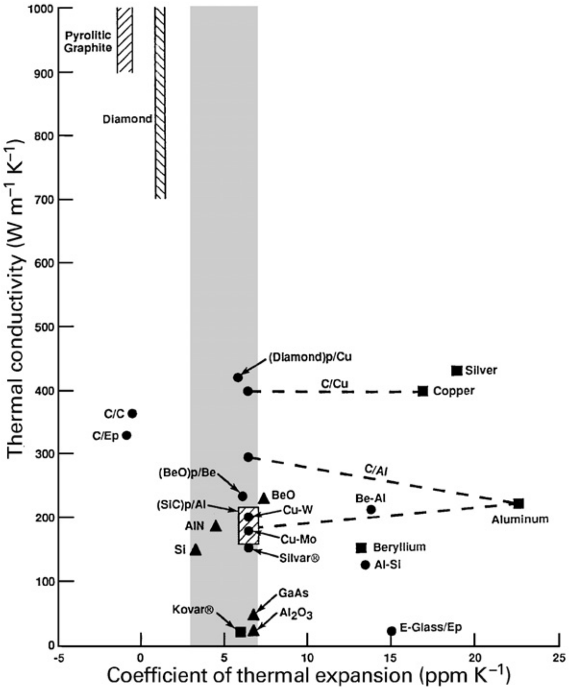


Figure 1
Thermal conductivity as a function of CTE for materials used in electronic packaging (reproduced by permission of John Wiley and Sons from "Mechanical Engineers' Handbook," 2nd edn., 1998).

Figure 1, which plots thermal conductivity as a function of CTE for a variety of materials used in electronics, illustrates the limitations of traditional packaging materials. In order to minimize thermal stresses in packaging, it is necessary to match the CTEs of semiconductors like silicon and gallium arsenide and ceramics used for substrates, such as alumina, beryllia, and aluminum nitride. These materials have CTEs in the range of about 3–7 ppmK^{-1} (shaded area). Aluminum and copper have good thermal conductivities, but, as Fig. 1 shows, CTEs that are much higher than desired.

Traditional materials with low CTEs include "Kovar," a nickel–iron alloy, and blends of copper and tungsten (Cu–W) or copper and molybdenum (Cu–Mo). Kovar has a low thermal conductivity and a density much higher than that of aluminum. To obtain low-CTE materials with higher thermal conductivities than Kovar, copper is blended with tungsten or molybdenum. Both liquid metal infiltration and powder metallurgy processes are used. As the two constituents are not truly alloyed, these materials can be considered composites. That is, they satisfy the definition of a composite material as two or more materials bonded together. For clarity, materials consisting of two distinct metals bonded together will be referred to as metal/metal composites. The term MMC is reserved for metals reinforced with carbon or ceramic fibers or particles. Table 1, which is based on data from Zweben (1998b), contains more detail about the properties of many of the materials in Fig. 1.

Note that although the thermal conductivities of Cu–W and Cu–Mo are much greater than that of

Table 1
Properties of selected electronic packaging materials.[a]

Reinforcement	Matrix or monolithic material	Thermal conductivity ($\mathrm{W\,m^{-1}\,K^{-1}}$)	CTE ($\mathrm{ppm\,K^{-1}}$)	Modulus (GPa)	Specific gravity
	Silicon	150	4.1		2.3
	Alumina	20	6.7	380	3.9
	Aluminum	120	23	69	2.7
	Copper	400	17	117	8.9
	Epoxy	1.7	54	3	1.2
	"Kovar"	17	5.9	131	8.3
Copper	Tungsten	167	6.5	248	16.6
Copper	Molybdenum	184	7.0	282	10.0
Beryllium	Aluminum	210	13.9	179	2.1
"Invar"	Silver	153	6.5	110	8.8
Carbon fibers (K1100)	Epoxy	300	−1.1	186	1.8
Carbon fibers (K1100)	Copper	400	6.5	158	7.2
Carbon fibers (K1100)	Aluminum	290	6.5	131	2.5
Discontinuous carbon fibers (K1100)	Polymer	20	4–7	30–50	1.6
Carbon fibers (K1100)	Carbon	350	−1.0	255	1.9
Silicon particles	Aluminum	126–160	6.5–13.5	100–130	2.5–2.6
SiC particles	Aluminum	170–220	6.2–7.3	225–265	3.0
Beryllia particles	Beryllium	240	6.1	330	2.6

Source: Zweben (1998b).

[a] CTEs and thermal conductivities for composites reinforced with continuous fibers are in-plane isotropic values. Fiber volume fractions for MMCs reinforced with continuous fibers are selected to achieve an in-plane CTE of 6.5 $\mathrm{ppmK^{-1}}$.

Kovar, they are in the range of aluminum alloys. Higher values, along with low CTE and density, are desired for many applications.

There are several materials in Fig. 1 that have CTEs in the desired range and significantly higher thermal conductivities than traditional materials: carbon fiber-reinforced copper (C/Cu), carbon fiber-reinforced aluminum (C/Al), diamond particle-reinforced copper ((Diamond)p/Cu), and beryllia particle-reinforced beryllium ((BeO)p/Be). There are a number of other materials that have high thermal conductivities, but which lie outside the shaded area. In addition to silver, aluminum, and copper, they include pyrolitic graphite, carbon/carbon (C/C) composites, carbon fiber-reinforced epoxy (C/Ep) and, of course, diamond.

To compare packaging materials in applications for which both thermal conductivity and density are important, a useful figure of merit is specific thermal conductivity, which is defined as thermal conductivity divided by density (for simplicity, we use specific gravity, which is dimensionless). Materials with high specific thermal conductivities are desirable in weight-critical applications. Figure 2 shows specific thermal conductivity as a function of CTE for packaging materials. As Fig. 2 shows, the specific thermal conductivities of traditional low-CTE packaging materials like Kovar, Cu–W, and Cu–Mo are very low. However, the low densities of many composites, combined with their high thermal conductivities, enhance their relative position. In fact, the specific thermal conductivities of some composites are more than an order of magnitude greater than those of traditional materials like Kovar, Cu–W, and Cu–Mo.

There is a rule of thumb to the effect that improving a critical property by an order of magnitude produces revolutionary change in a technology. It is not surprising, therefore, that composites are making significant inroads in electronic packaging.

A key advantage of some of the composites is that they can be used with processes that produce parts having the final desired shape (net shape processes) or which require only a small amount of machining (near-net shape processes). Minimizing machining enhances the cost effectiveness of these materials. This is particularly true for silicon carbide particle-reinforced aluminum, often called Al-SiC, which is one of the most important composite packaging materials.

In the following, the key types of reinforcements and composites used for thermal management, significant applications, and future directions are considered.

2. *Reinforcements*

There are several types of reinforcements used in composite materials: continuous fibers, discontinuous fibers, whiskers (elongated single crystals), and

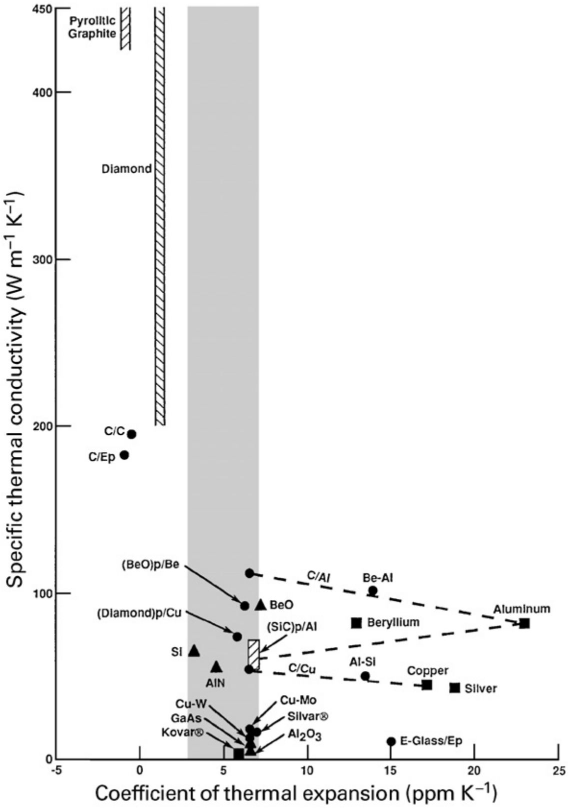


Figure 2
Specific thermal conductivity (thermal conductivity divided by specific gravity) as a function of thermal expansion for materials used in electronic packaging (reproduced by permission of John Wiley and Sons from "Mechanical Engineers' Handbook," 2nd edn., 1998).

particles. Flakes can be included in the last category. The most important reinforcements for obtaining high thermal conductivity are pitch-based carbon fibers, which are used in continuous and discontinuous forms, and two types of ceramic particles, silicon carbide (SiC) and beryllia (BeO). As a rule, continuous fibers are much more efficient than discontinuous fibers and particles for achieving high thermal conductivity, stiffness, and strength.

Composites reinforced with discontinuous fibers or particles may have processing advantages in some instances. A key characteristic of particle-reinforced materials is that they tend to be isotropic. That is, their properties are the same in all directions. This is not true of materials reinforced with continuous fibers, which are almost always anisotropic, which means that their properties vary with direction.

If a composite is reinforced with discontinuous fibers that are truly randomly oriented, its properties are statistically isotropic. However, virtually all processes produce some degree of preferred fiber orientation, and there are no strictly isotropic random-fiber composites. However, it is possible, and quite common, to make fiber-reinforced materials with properties that are approximately isotropic in a plane. These materials are called quasi-isotropic or pseudoisotropic materials.

A major advantage of using continuous fibers is that it is possible to vary directional properties significantly. For example, it is possible to produce composites with very high thermal conductivities, stiffnesses, and strengths in a particular direction by orienting a high percentage of fibers along this axis. In some instances, this can make fiber-reinforced composites more efficient than heat pipes for transporting heat over relatively short distances. In addition, solid-state methods of heat transfer are more reliable, and may be cheaper.

The most important types of commercial carbon fibers are made from polyacrylonitrile (PAN) and pitches derived from petroleum and coal. Figure 3 shows how thermal conductivity varies with electrical resistivity for selected metals and carbon fibers. As Fig. 3 shows, there are pitch-based carbon fibers, which are commercially available, with thermal conductivities more than twice that of copper. At the upper end of the spectrum are experimental fibers made by chemical vapor deposition (CVD), which have thermal conductivities in the region of $2000\,\mathrm{W m^{-1} K^{-1}}$, five times that of copper.

The most thermally conductive commercial carbon fiber, designated K1100, has a nominal axial thermal conductivity of $1100\,\mathrm{W m^{-1} K^{-1}}$ (Fleming *et al.* 1995).

Note that carbon fibers are generally strongly anisotropic, which means that their properties vary with direction. The radial thermal conductivity of thermally conductive pitch-based carbon fibers is at least an order of magnitude lower than the axial. In addition to their high thermal conductivities, pitch-based carbon fibers have other attractive attributes for packaging: elastic moduli as high as 965 GPa, more than an order of magnitude greater than that of aluminum, and very low densities. These reinforcements are used with polymer, metal, and carbon matrices to create light, stiff composites with extremely high thermal conductivities and low, controllable CTEs (Zweben and Schmidt 1989, Miller 1995, Fleming *et al.* 1995).

As previously noted, ceramic particles are used to increase the thermal conductivity and reduce the CTE of polymers. Although these particulate PMCs are significantly more thermally conductive than monolithic resins, their thermal conductivities are still an order of magnitude lower than that of aluminum. Since the early 1980s, however, ceramic particles have been added to metals to achieve the same objectives as for polymers, high thermal conductivity and low CTE, creating MMCs that are used in significant production applications. These composites also have

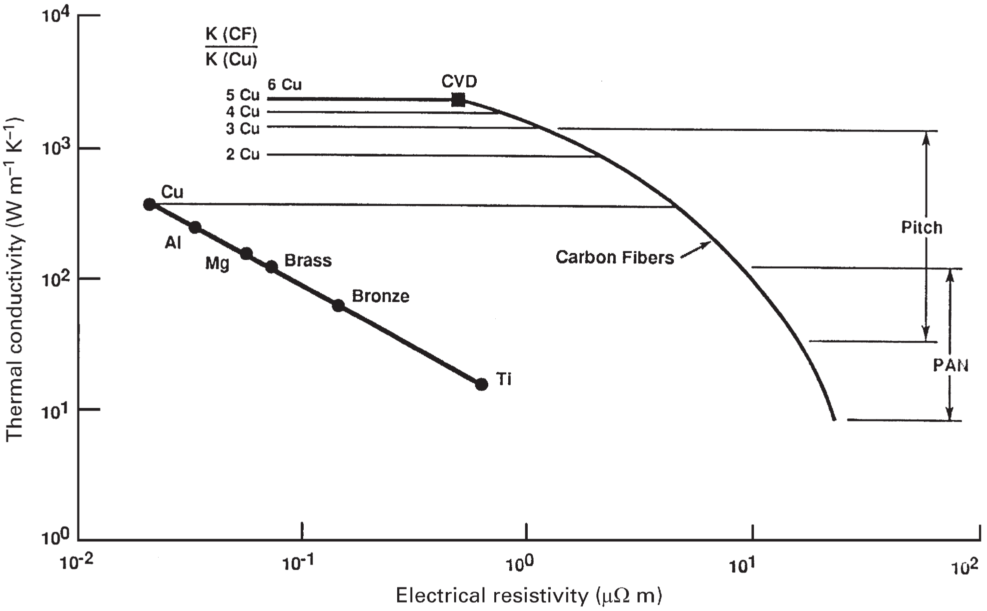


Figure 3
Thermal conductivity as a function of electrical resistivity for metals and carbon fibers (reproduced by permission of John Wiley and Sons from "Mechanical Engineers' Handbook," 2nd edn., 1998).

much lower densities than traditional low-CTE metallic materials used in packaging (Thaw *et al.* 1987, Zweben 1998a, Occhionero *et al.* 1999).

As mentioned above, the two key materials used as particulate reinforcements are silicon carbide and beryllia. Diamond particles also have been used to reinforce metals on an experimental basis (Zweben 1998). Because of the difficulty of measuring the thermal conductivity of particles, it is common to assume that they have values similar to those for bulk materials. However, there are considerable differences in published thermal conductivities for ceramics. In general, transport properties, like thermal conductivity, are more difficult to measure than bulk properties such as modulus.

3. Composite Heat Sink Materials

As discussed above, there are three key classes of composite materials used in electronic packaging thermal management: PMCs, MMCs, and CCCs.

3.1 PMCs

Polymers reinforced with continuous pitch-based carbon fibers have the highest thermal conductivity of all commercial PMCs. Table 1 gives the in-plane properties of quasi-isotropic epoxy-matrix composites reinforced with continuous K1100 fibers. Composite properties depend strongly on reinforcement content. The fiber volume fraction here is 0.60. The in-plane thermal conductivity and elastic modulus of this composite are much higher than those of aluminum, and its density is much lower. The specific thermal conductivity of this composite is greater than that of copper. Note that the in-plane CTE is negative. In fact, most commercial carbon fibers have negative CTEs. Although the CTE of this composite does not fall in the desired range of 3–7 ppm K^{-1}, it is much closer than those of aluminum and copper.

Note that a K1100 carbon fiber-reinforced PMC with a fiber volume fraction of 0.6 having all fibers oriented in the same direction, which is called a unidirectional composite, would have a nominal axial thermal conductivity of 660 W m^{-1} K^{-1}, over 150% greater than that of copper (about 400 W m^{-1} K^{-1}).

Although continuous fibers offer the best properties, PMCs reinforced with discontinuous fibers can be used with relatively inexpensive net shape processes such as injection molding (Fleming *et al.* 1995). Table 1 gives representative properties of a composite with discontinuous K1100 fibers. Note that both thermal conductivity and elastic modulus are much lower than for composites reinforced with continuous fibers.

3.2 MMCs

The key electronic packaging MMCs are silicon carbide particle-reinforced aluminum (often called Al-SiC), beryllia particle-reinforced beryllium, and carbon fiber-reinforced aluminum (Zweben 1998a). Table 1 lists properties of these materials.

A number of other MMCs have been developed, including beryllium–aluminum, "Invar"–silver ("Silvar"), and "Invar"–copper ("Cuvar"). A commercially available material consisting of 60% aluminum and 40% silicon also has been used for packaging. This material can be considered a type of particulate MMC, because of the low mutual solubility of silicon and aluminum at room temperature. A family of Al–Si materials with higher levels of silicon has been made by a powder spray process (Zweben 1998a). One formulation has a CTE in the desired range.

Perhaps the most successful of the newer MMCs is Al-SiC, which was first used in packaging in the 1980s (Thaw *et al.* 1987). Figure 4 shows how the CTE of Al-SiC varies with particle content. For particle volume fractions in the region of 0.7 it is possible to match the CTE of silicon. That, and its low density and net shape fabrication capability, have led to rapidly increasing use of this material (Occhionero *et al.* 1999, Zweben 1992, 1998a). Carbon fiber-reinforced aluminum also has been used in a limited number of highly specialized applications, such as spacecraft heat sinks and radiator panels.

Beryllia particle-reinforced beryllium also is used in commercial applications, although not as widely as Al-SiC because of its higher cost (Zweben 1998a).

Like common laminated metals used in electronic packaging, such as copper–"Invar"–copper and copper–molybdenum–copper, MMCs reinforced with continuous fibers may display significant hysteresis when subjected to large thermal excursions or mechanical stresses (Zweben 1989).

Other MMCs with high thermal conductivities are made by infiltrating porous CCCs with aluminum or copper. For copper matrix composites reinforced with highly oriented, thermally conductive carbonaceous fibrous materials, such as "ThermalGraph," axial thermal conductivities as high as 890 $\mathrm{W m^{-1} K^{-1}}$ have been reported (Miller 1995).

3.3 CCCs

There are a large number of materials that fall in this category, including "ThermalGraph," mentioned above. CCCs made from a pitch-based carbon fiber, P120, which has a nominal axial thermal conductivity of about 600 $\mathrm{W m^{-1} K^{-1}}$, are reported to have an in-plane isotropic thermal conductivity of 250 $\mathrm{W m^{-1} K^{-1}}$

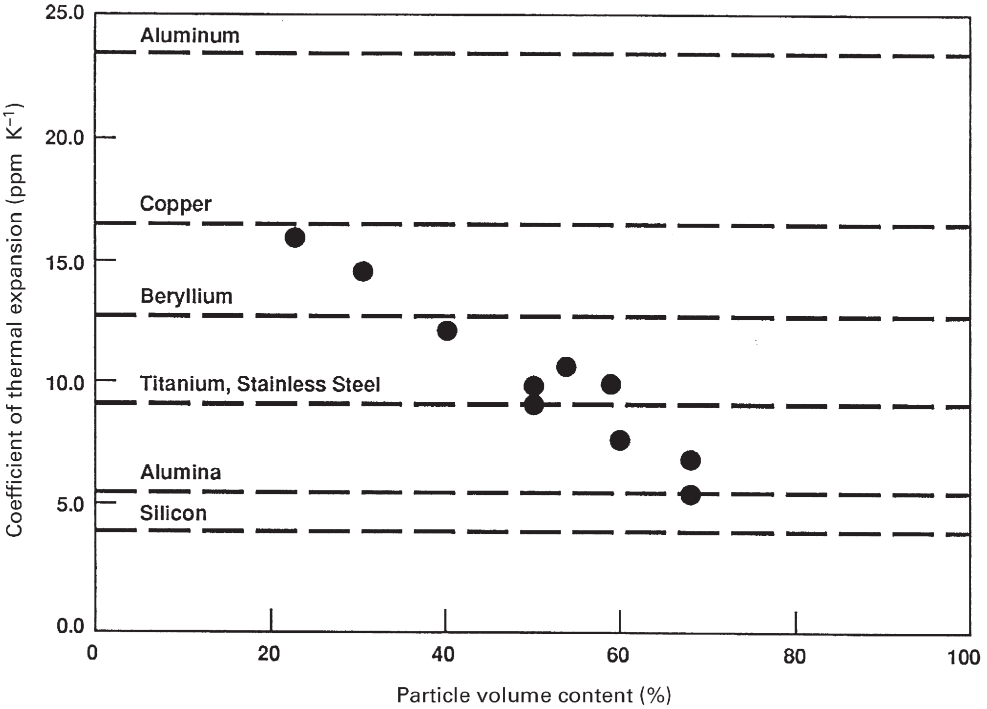


Figure 4
CTE as a function of particle content for silicon carbide particle-reinforced aluminum (reproduced by permission of John Wiley and Sons from "Mechanical Engineers' Handbook," 2nd edn., 1998).

(Miller 1995). This, combined with a specific gravity of only 1.85, makes CCCs very attractive thermal control materials for weight-sensitive applications (Shih 1994).

4. Pyrolitic Graphite

As shown in Figs. 1 and 2, some forms of pyrolitic graphite, a monolithic material, can have very high in-plane thermal conductivity (values of $1200\,\mathrm{W m^{-1} K^{-1}}$ are reported), combined with a relatively low specific gravity of 2.5 (Montesano 1996). However, this material is highly anisotropic, and the through-thickness conductivity is only about $25\,\mathrm{W m^{-1} K^{-1}}$. Pyrolitic graphite is used in a form in which it is encapsulated with materials such as aluminum, copper, beryllium, "Kovar," and two composites, Al-SiC and C/Ep.

5. Applications

Advanced materials, especially composites, are now being used in a significant and increasing number of commercial and aerospace electronic packaging thermal management applications. Components include carriers; complex hermetic microwave packages (also called modules); power semiconductor packages; solid and liquid-cooled, flow-through PCB heat sinks; heat sink lids for central processor units; thermal straps; enclosures; and support structures. Pyrolitic graphite is also used in heat sink applications.

The components cited above are used in numerous commercial and aerospace products, including electric and hybrid vehicles; motor controllers; cellular telephone ground stations; laptop computers; and aircraft and spacecraft microwave and power supply subsystems. In addition, composites are key materials in developmental integrated multifunctional structural/electronic systems.

6. Summary

It is a virtual certainty that packaging density will continue to increase, and with it the need for continuing improvements in thermal management. This will require a variety of solutions, including use of heat pipes and active fluid cooling. However, it is unlikely that the latter two approaches will solve all thermal management problems. Further, a passive solution, such as use of materials with improved thermal conductivities, simplifies thermal management, improves reliability, and may reduce cost. In addition, use of active cooling and heat pipes does not eliminate the issue of thermal stresses resulting from thermal expansion mismatch.

Another important trend is the continuing competition to reduce the weight of transportable systems, such as laptop computers and hand-held devices such as personal digital assistants and cellular telephones. In addition to these commercial products, the drive to reduce weight in electronics used in aircraft, spacecraft, and even shipboard applications is likely to continue.

Based on these facts, it seems likely that there will continue to be a need for improved materials having high thermal conductivities and low, controllable CTEs. In addition, there is a need for low-density materials in many applications.

It is clear that we are in the early stages of an important new materials technology. In the future, it is reasonable to expect continuing development of improved and new reinforcements, matrix materials, material systems, and manufacturing processes. Based on past experience, it is likely that increasing material production volumes and improved processes will reduce costs, making composites and other advanced packaging materials more competitive. For example, carbon fibers sold for US\$500 $\mathrm{kg^{-1}}$ when they were first introduced commercially in the early 1970s. It is now possible to buy commercial structural-grade fibers for as little as US\$4 $\mathrm{kg^{-1}}$. Further price reductions appear likely as production volume increases.

See also: Electronic Packaging Design: Fundamentals#; Electronic Packaging Materials: Properties and Selection#

Bibliography

Chung D D L 1995 Overview of materials for electronic packaging. In: Chung D D L (ed.) *Materials for Electronic Packaging*. Butterworth-Heinemann, Oxford

DiNardo M, Kreitz J, Zweben C 1990 Lightweight composite electronics enclosure for high stress and temperature environments. *Proc. 35th Int. SAMPE Symp. Exhibition*. Anaheim, CA

Fleming T F, Levan C D, Riley W C 1995 Applications for ultra-high thermal conductivity fibers. *Proc. Int. Electronic Packaging Conf.* International Electronic Packaging Society, Wheaton, IL, pp. 493–503

Kelly A 1988 *Concise Encyclopedia of Composite Materials*. Pergamon, Oxford

Miller J 1995 "ThermalGraph" product forms for use in space, electronics and satellite applications. *Proc. Semin. Pitch Fiber Composites for Space Structures*. Long Beach, CA

Montesano M J 1996 New material for thermal management has four times thermal conductivity of copper. *Mater. Technol.* **11** (3), 87–91

Occhionero M, Hay R A, Adams R W, Fennessy K P 1999 Cost-effective manufacturing of aluminum silicon carbide (AlSiC) electronic packages. *Proc. IMAPS Advanced Packaging Materials Symp.* Braselton, GA

Shih W T 1994 Carbon–carbon (C–C) composites for thermal plane applications.*Proc. 7th Int. SAMPE Electronics Conf.*

Thaw C, Zemany J, Zweben C 1987 Metal matrix composite microwave packaging components. *SAMPE J.* **23** (6)

Zweben C 1989 Mechanical and thermal properties of silicon carbide particle-reinforced aluminum. In: Kennedy L M, Moeller H H, Johnson W S (eds.) *Thermal and Mechanical Behavior of Metal Matrix and Ceramic Matrix Composites*, ASTM STP 1080. American Society for Testing and Materials, Philadelphia

Zweben C 1992 Overview of metal matrix composites for electronic packaging and thermal management. *JOM* (July), 15–23

Zweben C 1995 The future of advanced composite electronic packaging. In: Chung D D L (ed.) *Materials for Electronic Packaging*. Butterworth-Heinemann, Oxford

Zweben C 1998a Advances in composite materials for thermal management in electronic packaging. *JOM* (June), 47–51

Zweben C 1998b Composite materials and mechanical design. In: Kutz M (ed.) *Mechanical Engineers' Handbook*, 2nd edn. Wiley, New York

Zweben C, Schmidt K A 1989 Advanced composite packaging materials. In: *Electronic Materials Handbook*. ASM International, Materials Park, OH

C. Zweben
Devon, Pennsylvania, USA

F

Fatigue of Particle Reinforced Materials

The combination of a light metal (such as aluminum or magnesium) and a high-modulus reinforcement (typically a ceramic such as SiC or Al_2O_3) results in improved strength and stiffness over conventional light metallic alloys, while providing weight savings and allowing higher operating temperatures (Clyne and Withers 1993, Chawla 1997). Much of the driving force for engineering such metal matrix composites (MMCs) has been that monolithic lightweight alloys have inadequate fatigue resistance for many demanding applications. The use of a high-stiffness ceramic reinforcement in particulate form, such as SiC, can result in a substantial increase in fatigue resistance while maintaining cost at an acceptable level (Chawla and Shen 2001). This is an important design criterion in many applications, e.g., in automotive components. An example of an application where a lightweight, fatigue-resistant component is required is the connecting rod in passenger cars. The fatigue resistance of particulate MMCs depends on a variety of factors, including reinforcement particle volume fraction, particle size, matrix microstructure, the presence of inclusions or defects that arise from processing, and testing environment. The effect these factors have on the fatigue behavior of particle reinforced MMCs is summarized in this article (see also *Composite Materials, Overview*#).

1. *Strengthening in MMCs*

In general, strengthening in MMCs can be classified into direct and indirect strengthening. When a load is applied to a particulate composite, the load is transferred from the soft metallic matrix to the stiff ceramic reinforcement. This classical composite strengthening is an example of direct strengthening. The strength of the composite may also be enhanced by the influence reinforcement may have on matrix microstructure or deformation mode. For example, on cooling, there is a thermal mismatch between the low-expansion ceramic and the high-expansion metal. This thermal mismatch between the reinforcement and the matrix can result in the introduction of additional dislocations that strengthen the metal matrix. This is termed "indirect" strengthening, because the addition of the reinforcement indirectly strengthens the matrix of the composite. The indirect effect may be enhanced in precipitation hardened aluminum alloys where precipitate nucleation during aging takes place at dislocations. Thus, the intrinsic properties and microstructure of the matrix material in the composite may be significantly different from those of the unreinforced alloy.

2. *Effect of Volume Fraction and Particle Size*

Increasing the volume fraction and decreasing the particle size of the reinforcement both result in enhanced fatigue resistance, from a combination of direct and indirect strengthening. In the composite, most of the load is carried by the high-modulus, high-strength reinforcement, so for a given stress the composite undergoes a lower strain than the unreinforced alloy. Thus, the fatigue lives of particle-reinforced MMCs are generally longer than those of unreinforced metals. This is shown for a 2080 (Al–Cu–Mg) matrix composite reinforced with SiC particles in Fig. 1. These improvements are most pronounced at lower stresses, in the high cycle fatigue regime, while at high stress the differences between reinforced and unreinforced materials are reduced. With decreasing particle size, for a given reinforcement volume fraction, the reinforcement interparticle spacing decreases, resulting in more barriers for the reversible slip motion that takes place during fatigue, and a decrease in strain localization by cyclic slip refinement. Above a critical particle size, reinforcement fracture is predominant and contributes to premature fatigue life, because of the increased propensity for particle cracking as the particle size increases. Narrowing of the particle size range distribution also results in higher fatigue life, particularly when eliminating larger particles that are more prone to cracking.

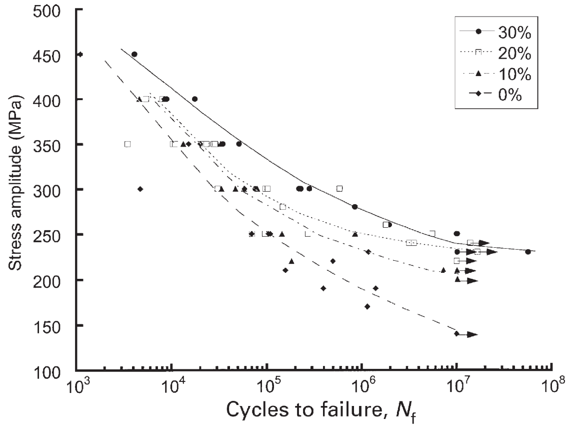


Figure 1
Effect of reinforcement volume fraction (5 μm SiC particle size) on fatigue life of a $2080/SiC_p$ composite (Chawla *et al.* 1998). Increasing volume fraction of the reinforcement results in enhanced fatigue resistance.

---

* Cross references marked by an asterisk are included in this volume.

# Cross references marked by a hash can be found by consulting the Encyclopedia of Materials: Science and Technology.

3. *Effect of Matrix Microstructure*

In addition to particle reinforcement, the matrix microstructure also significantly influences the fatigue behavior of the composite. Factors affecting the matrix microstructure include size, shape, and spacing of precipitates, grain size, and nonreinforcement dispersoids or inclusions (such as iron-rich inclusions that are formed during processing). The effect of grain size has not been explicitly examined, but composites follow the same trend as monolithic materials, i.e., for a given matrix alloy composition and volume fraction of reinforcement, finer grain sizes generally result in improved properties.

Contrary to conventional wisdom for monolithic materials, in MMCs high matrix yield and ultimate tensile strength do not necessarily reflect high fatigue strength (defined here as fatigue runout at 10^7 cycles). There is no significant difference in fatigue behavior between naturally aged and artificially aged MMCs, even though the naturally aged material typically has a much lower yield and ultimate strength. In age-hardenable matrix composites, thermal and thermomechanical treatments have a significant effect on the matrix microstructure and the resulting properties of the composite. A thermomechanical treatment (T8) typically produces a fine and homogeneous distribution of precipitates, while a thermal treatment only (T6) generally results in coarser and inhomogeneously distributed precipitates. Because of finer and more closely spaced precipitates, the composite that undergoes a T8 treatment usually exhibits higher yield strengths than that of T6 materials. Despite their lower yield strength, however, the T6 matrix composites have been shown to exhibit higher fatigue resistance than T8 matrix composites. This is shown in Fig. 2 for 2080/SiC_p-T6 and 2080/SiC-$T8_p$ of the same reinforcement volume fraction and particle size. The contrasting behavior between monotonic and cyclic loading in this composite can be attributed to the strong influence of the presence, stability, and morphology of S′ (Al_2CuMg) precipitates in the matrix of the composite. In fatigue, failure processes are affected by a variety of microstructural factors, which include resistance to dislocation motion and pileup at precipitates and/or reinforcement particles and cracking at slip bands. Throughout the fatigue process, dislocations either cut precipitates or loop around them during cyclic deformation, depending on precipitate coherency and size. In T8 materials, the precipitates are fine enough such that it is believed that the precipitates are cut by dislocations and localized slip takes place, which reduces the strengthening effect of precipitates, thereby impairing the fatigue strength. In T6 materials, the larger precipitate size allows them to retain their precipitate structure and strength during fatigue.

Overaging heat treatments also modify the matrix microstructure, resulting in coarsening of the precipitate structure, while retaining a homogeneous precipitate distribution, which directly influences fatigue life. Figure 3 shows the coarsening and increase in precipitate spacing in 2080/SiC_p-T8, from the strained and peak-aged condition, to a significantly overaged condition. Increasing precipitate spacing decreases both fatigue strength and fatigue lifetime (Fig. 4). This is to be expected since coarser precipitates mean a larger interprecipitate spacing and easier bypass of dislocations. For the composites subjected to higher overaging temperatures, the yield strength and fatigue strength also decrease with increasing precipitate spacing.

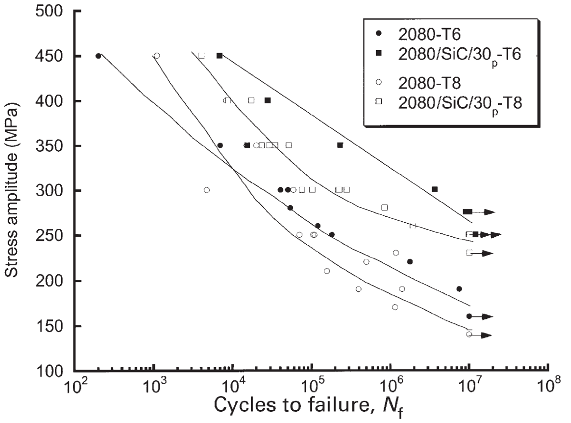


Figure 2
Comparison of fatigue life of unreinforced and 2080/SiC_p composites in peak-aged (T6) and rolled and peak-aged (T8) conditions. The T8 matrix microstructure consists of relatively fine and closely spaced precipitates that produce a material with high yield strength (Chawla *et al.* 2000b). The T6 matrix materials have relatively larger and more inhomogeneous distribution of precipitates, resulting in lower yield strength.

Processing-related defects in the form of intermetallic inclusions or particle clusters are also part of the matrix microstructure and play a role in fatigue strength, particularly in powder metallurgy processed materials. These defects act as stress concentrators that increase the local stress intensity in the material and promote easy crack nucleation. Crack initiation during fatigue takes place at these defects, which are typically located at the surface of the specimen. This is because inclusions at the surface are more highly stressed than inclusions completely within the matrix, so a higher stress concentration and, thus, higher probability for crack initiation is present at the surface. For a given inclusion size, the stress concentration in a composite, where the inclusion is surrounded by high-stiffness reinforcement particles, is lower than in an unreinforced alloy. Since more of the load is being "shared" by the high-stiffness SiC particles in the composite, an inclusion in the com-

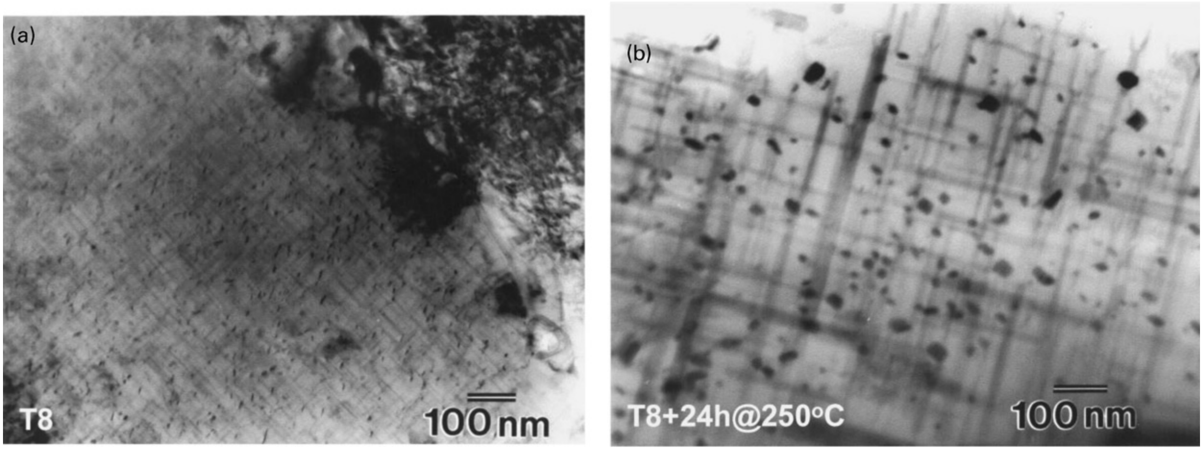


Figure 3
Precipitate coarsening and increase in precipitate spacing in MMCs overaged at 225 °C for 24 h (Chawla *et al.* 2000b).

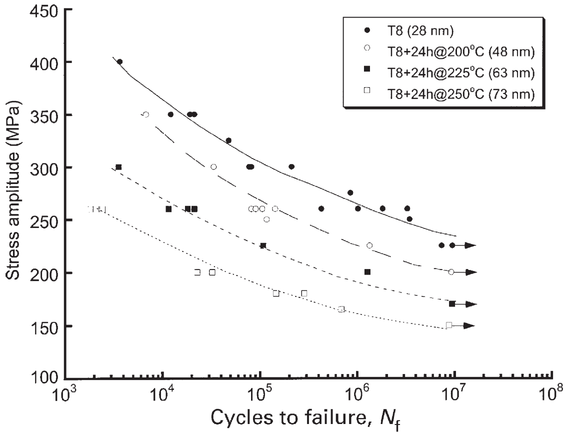


Figure 4
Effect of overaging (increase in precipitate size and spacing) on fatigue life (Chawla *et al.* 2000b). Precipitate spacings are given in parentheses.

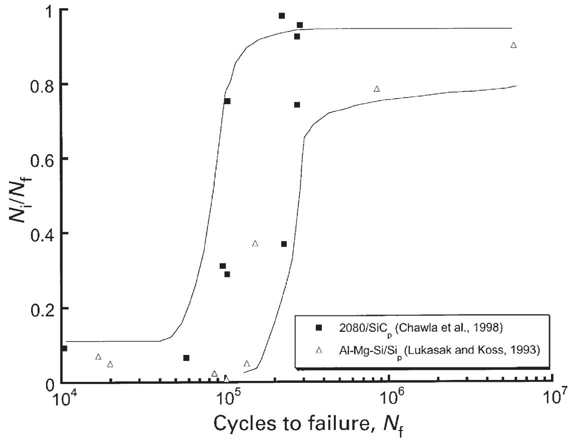


Figure 5
Ratio of cycles to initiation to cycles to failure (N_i/N_f) vs. fatigue life for two different composites systems (Lukasak and Koss 1993, Chawla *et al.* 1998). Notice the transition from crack propagation being dominant in the low cycle regime, compared to most of the fatigue life being spent initiating a crack in the high cycle fatigue regime.

posite will be subjected to lower stress than a similar inclusion in the unreinforced alloy. In extruded materials, the overall size of inclusions is also lower in composites since the ceramic reinforcement particles break the brittle inclusions into smaller sizes during extrusion.

It is interesting to note that in the low-cycle regime, cracks seem to originate relatively early in fatigue life (around 10% of total life). In the high-cycle regime, however, crack initiation can occur quite late (after about 70–90% of the life of the specimen). Figure 5 shows data for two different composites systems that indicates the transition from crack propagation being dominant in the low-cycle regime, compared to most of the fatigue life being spent initiating a crack in the high cycle fatigue regime. While crack growth is relatively unimpeded in unreinforced materials (Fig. 6(a)), crack growth is hindered by mechanisms such as crack deflection and crack trapping in the composite (Fig. 6(b)).

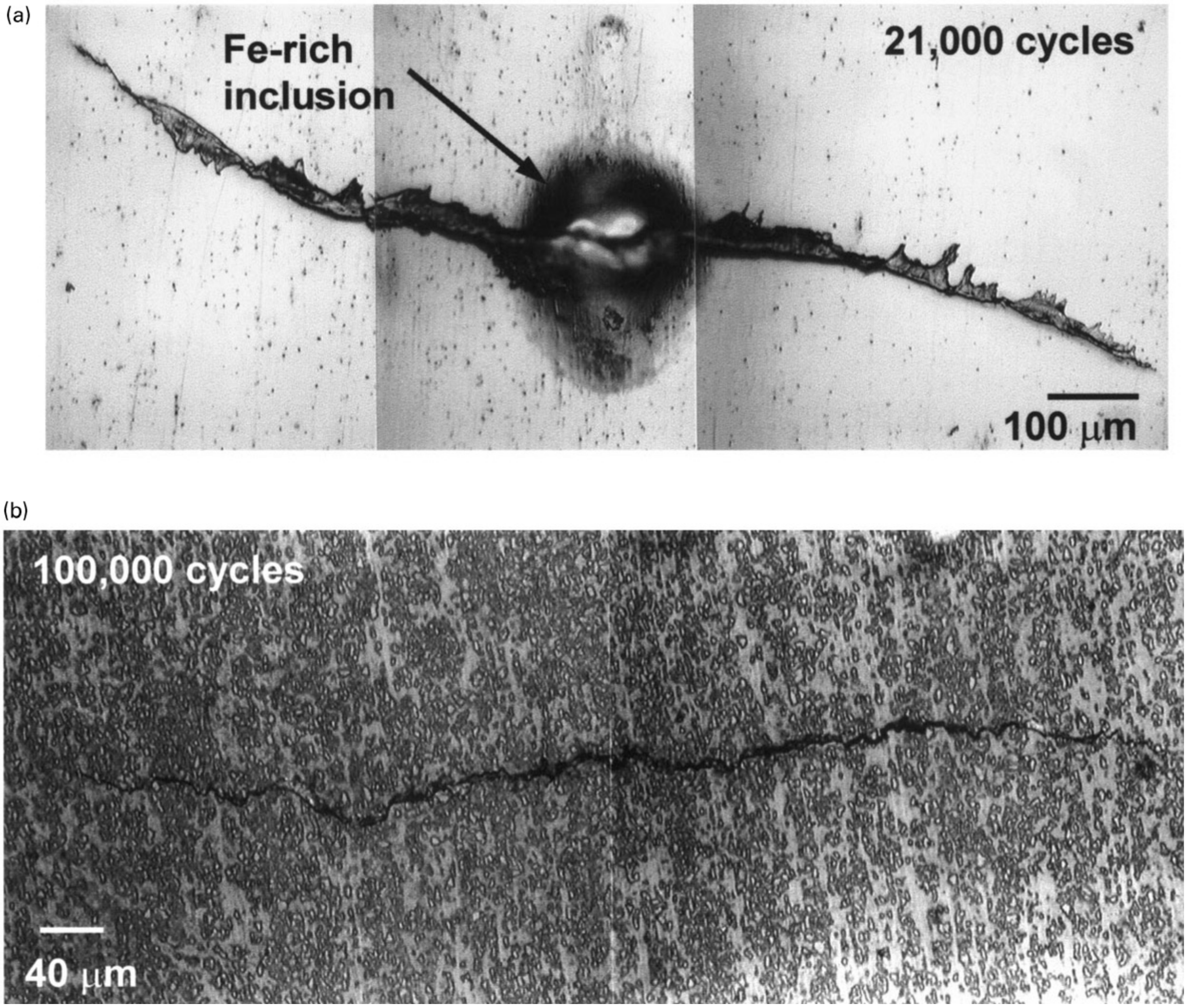


Figure 6
Fatigue crack initiation and growth in (a) unreinforced 2080 Al and (b) 2080/SiC$_p$ composite (Chawla *et al.* 2000a). Notice that crack growth is hindered by mechanisms such as crack deflection and crack trapping in the composite.

4. *Effect of Fatigue Testing at Elevated Temperatures*

As described above, several MMC applications, such as the connecting rod in passenger cars, require fatigue resistance at elevated temperatures (of the order of 150–175 °C). Figure 7 shows the elevated temperature fatigue performance of 2080/SiC$_p$-T6 composites in comparison with the room temperature data shown in Fig. 1. The most significant loss of fatigue strength occurs between 25 °C and 150 °C. The incremental increase in temperature to 170 °C results in a modest decrease in fatigue strength. In applications where elevated temperature fatigue resistance is a criterion, the temperature at which the composite is aged becomes very important. At slightly higher temperatures than the aging temperature, a severe decrease in fatigue strength may take place because of significant matrix overaging. The fatigue strength in the composites seems to be directly proportional to the strength of the matrix, as shown by the fatigue ratio for an MMC tested at 25 °C and 150 °C (Table 1). Table 1 also shows that the decrease in fatigue strength owing to temperature is significantly higher than the decrease in yield strength. This may be caused by changes in matrix microstructure and decrease in matrix strength from a combination of long-term exposure and cyclic stress at elevated temperatures. Note that

Table 1
Effect of reinforcement volume fraction on fatigue properties of an MMC.

Material	Approximate fatigue strength (10^7 cycles), σ_{fat} (MPa)	Yield strength (0.2% offset), $\sigma_{0.2}$ (MPa)	Fatigue ratio ($\sigma_{\text{fat}}/\sigma_{0.2}$)
2080/SiC/30_p-T6, 25 °C	275	529	0.52
2080/SiC/20_p-T6, 25 °C	250	467	0.54
2080/SiC/30_p-T6, 150 °C	210	470	0.45
2080/SiC/20_p-T6, 150 °C	200	435	0.46

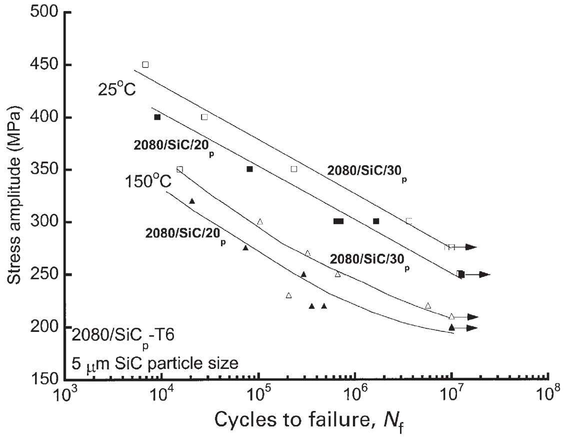


Figure 7
Elevated temperature fatigue performance of an MMC in comparison with the room temperature data shown in Fig. 1 (Chawla *et al.* 1999). The 2080/SiC/30_p material with 5 μm particle size meets the stringent fatigue requirements for a connecting rod.

the composite meets the stringent connecting rod target of 210 MPa fatigue strength at 150 °C.

5. *Summary*

Some observations can be made about the key issues in designing MMCs for fatigue resistance. A high volume fraction of the particulate reinforcement is needed to obtain the desired elastic modulus of the composite for fatigue resistance, while relatively fine particle sizes also improve fatigue strength. The precipitate size, structure, and spacing in the matrix are very important in determining fatigue properties. Fine, closely spaced precipitates that provide high yield and tensile strength may be sheared by dislocations during the fatigue process. Relatively larger precipitates may retain their integrity during the fatigue process and provide a higher fatigue life, although very large, incoherent precipitates, as observed in overaged microstructures, are not conducive to fatigue resistance. The minimization of the size and number of nonreinforcement inclusions or defects in the matrix also increases fatigue life. Thermal stability of the matrix is important, since a thermally stable matrix is required for long-duration applications such as the connecting rod in automotive engines. Finally, the quality of the fabrication process of the composites is extremely critical to obtain a composite with large volume fraction and fine particle size reinforcement, and a thermally stable microstructure with a minimum of defects. More importantly, low-cost processing routes are needed to implement these materials in high-volume applications.

Bibliography

Chawla K K 1997 *Composite Materials—Science and Engineering*, 2nd edn. Springer, New York, p. 102

Chawla N, Shen Y-L 2001 Mechanical behavior of particle reinforced metal matrix composites–a review. *Adv. Eng. Mater.* **3**

Chawla N, Andres C, Davis L C, Jones J W, Allison J E 2000a The interactive role of inclusions and SiC reinforcement on the high cycle fatigue resistance of particle reinforced metal matrix composites. *Metall. Mater. Trans.* **31A**, 951–7

Chawla N, Andres C, Jones J W, Allison J E 1998 Effect of SiC volume fraction and particle size on the fatigue resistance of a 2080-Al/SiC_p composite. *Metall. Mater. Trans.* **29A**, 2843–54

Chawla N, Habel U, Shen Y-L, Andres C, Jones J W, Allison J E 2000b Effect of matrix microstructure on the fatigue behaviour of SiC particle reinforced 208° Al matrix composites. *Metall. Mater. Trans.* **31A**, 531–40

Chawla N, Jones J W, Allison J E 1999 Elevated temperature tensile and fatigue behavior of particle reinforced metal matrix composites. In: Wu X R, Wang Z G (eds.) Fatigue *'99*. EMAS/HEP

Clyne T W, Withers P J 1993 *An Introduction to Metal Matrix Composites*. Cambridge University Press, Cambridge, p. 1

Lukasak D A, Koss D A 1993 Microstructural influences on fatigue crack initiation in a model particulate-reinforced aluminum-alloy MMC. *Composites* **24**, 262–9

N. Chawla[a] and J. E. Allison[b]
[a] *Arizona State University, Tempe, Arizona, USA*
[b] *Ford Motor Company, Dearborn, Michigan, USA*

Fiber Metal Laminates, Fatigue of

Fiber–metal laminates were developed as a family of hybrid materials that consist of bonded thin metal sheets and fibers embedded in adhesive, i.e., rubber-toughened epoxy (Fig. 1). This laminated layout creates a material with excellent fatigue, impact, and damage tolerance characteristics, and a low density (Vogelesang and Gunnink 1986, Vlot *et al*. 1999). The bondlines act as barriers against corrosion of the metallic sheets, and the metal layers protect the fiber/epoxy layers from picking up moisture. The laminate has an inherent high burn-through resistance as well as good damping and insulation properties. The material can be produced as semifinished sheet material; like monolithic aluminum, these sheets can be machined and formed into products. The laminate can also be cured in an autoclave into a complete structure: large, curved panels with co-cured doublers and stiffening elements. The so-called "splicing concept" makes a larger panel size possible compared to conventional aluminum structures, with consequent reduced assembly costs. The possible 20% weight reduction was the prime reason behind the development of these materials for aircraft structures.

1. Background and Different Types

The idea of fiber–metal laminates originated in the 1950s at the Dutch Fokker aircraft company. Instead of integral machined structures, it was chosen to adapt the local strength of laminated bonded structures, mostly for cost reasons. It turned out that as an added benefit, owing to the bending present in most structures, fatigue cracks in laminated aluminum-based structures remain mainly limited to the outer layer for a large portion of their life, while the remainder of the layers restrain crack opening. In the late 1970s the first tests were performed with fibers that were meant to reinforce the bondline; these attempts were not successful. In the early 1980s, however, fiber–metal laminates were developed as an optimized, laminated, semifinished sheet product (Vogelesang *et al*. 1981). This was achieved by creating an optimized, carefully balanced mix of thin aluminum layers combined with fiber/epoxy layers. The optimal metal fraction relative to the fiber/epoxy content appeared to be crucial for an adequate fatigue performance. The first generation of laminates was based on aramid fibers and was named ARALL (aramid-reinforced aluminum laminate). These laminates had unidirectional fiber layers and were developed primarily for aircraft wing applications. ARALL reached maturity when its promising characteristics were finally demonstrated by testing a full-scale Fokker 50 wing panel in the period 1984–87, which proved that a weight saving of 20% was possible (Jongebreur *et al*. 1992). ARALL was selected for the cargo doors for the US Air Force's C-17 aircraft.

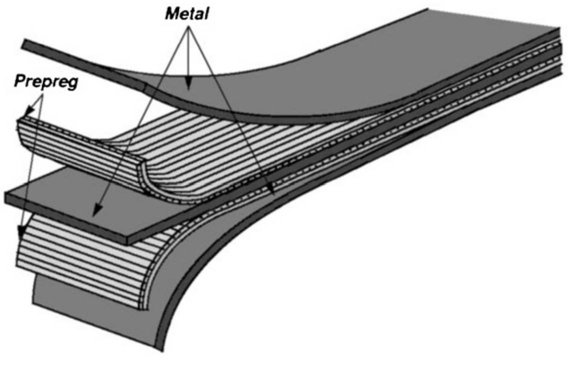


Figure 1
Typical fiber–metal laminate lay-up.

In 1987 the second generation of laminates was born with the name GLARE (glass reinforced), which was based on high-strength glass fibers (Roebroeks 1991). Whereas ARALL had unidirectional fiber layers because it was optimized for wing structures with a uniaxial stress state, GLARE was developed both in unidirectional and cross-plied variants. GLARE had to be developed for fuselage applications because fiber failure was found to occur in ARALL under realistic fuselage skin fatigue loading (i.e., a constant amplitude loading between zero and maximum load) owing to the bad compression properties of the aramid fiber. Because internal compressive stresses are present after curing the laminate at 120 °C owing to a mismatch of the coefficients of thermal expansion of the constituents, cyclic compression of the aramid fibers will occur even if the minimum stress on the laminate is zero. The glass fibers in GLARE do not show failure under realistic fuselage fatigue loads. GLARE soon appeared to be also superior for applications other than fuselage skins. The technology of GLARE was extended in close cooperation with Deutsche Airbus and extensively demonstrated in 1988–89 in Germany by testing a full-scale Airbus A330/340 fuselage barrel.

GLARE is manufactured by laminating the fiber/epoxy prepreg and aluminum layers, after which the laminate is cured for two hours in an autoclave at a temperature of 120 °C under an air pressure of 6 bar. After this curing, the laminate has an internal stress system with tension in the aluminum layers and compression in the fibers. This stress state can be reversed by plastically stretching the laminate. This causes a permanent elongation of the aluminum layers, whereas the fiber layers deform elastically. Therefore, after this operation the fibers impart a compressive stress on the elongated aluminum layers. Several GLARE variants are commercially available (see Table 1). Every grade can be built up in many different thicknesses, e.g., the lay-up of a GLARE 3

Table 1
Available GLARE variants.

	Aluminum layer		Fiber layer	
Variant	Alloy	Thickness per layer (mm)	Orientation	Thickness per layer (mm)
GLARE 1	7475-T761	0.3–0.4	Unidirectional	0.25
GLARE 2	2024-T3	0.2–0.5	Unidirectional	0.25
GLARE 3	2024-T3	0.2–0.5	Cross-ply (50%–50%)	0.25
GLARE 4	2024-T3	0.2–0.5	Cross-ply (67%–33%)	0.375
GLARE 5	2024-T3	0.2–0.5	Cross-ply (50%–50%)	0.5

variant with a so-called "3/2 lay-up" as shown in Fig. 1 is:

[2024-T3/0°glass/90°glass/2024-T3/90°glass/
0°glass/2024-T3]

with the 0° direction as the L-direction (rolling direction of the aluminum layers).

2. *Fatigue Behavior in Experiments*

The key to the excellent fatigue performance is the so-called "crack bridging mechanism" (Fig. 2). The crack opening in the aluminum layers is restrained by the fibers. In other words, the stress intensity at the crack tip in the aluminum layers is reduced by effective crack closing stresses induced by the fibers on the crack flanks. The fibers do not fail during crack opening because cyclic shear stresses in the adhesive create a controlled delamination growth around the crack under fatigue loading. This balance of delamination and crack growth ensures a distribution of strain created by crack opening over a larger length of the fibers, which prevents fiber failure.

Figure 3 compares the fatigue performance of fiber–metal laminates and aluminum 2024-T3 under simulated fuselage loading from a central saw cut. While the crack growth rate of the monolithic material increases rapidly with increasing crack length, the laminates exhibit their characteristic, very slow crack growth behavior. Tests indicate that over a large range of crack lengths the crack growth rate is constant. Under realistic loading conditions, fiber–metal laminates exhibit crack growth rates of the order of 10–100 times slower than their monolithic aluminum constituents.

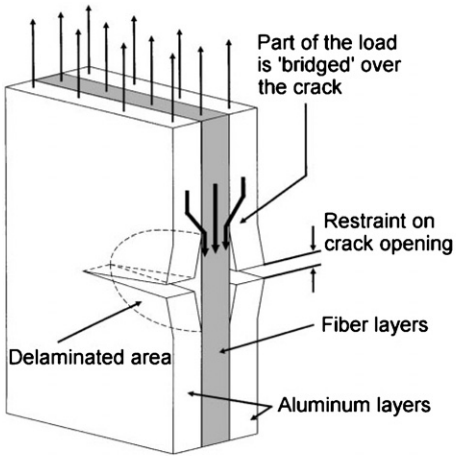


Figure 2
Crack bridging mechanism.

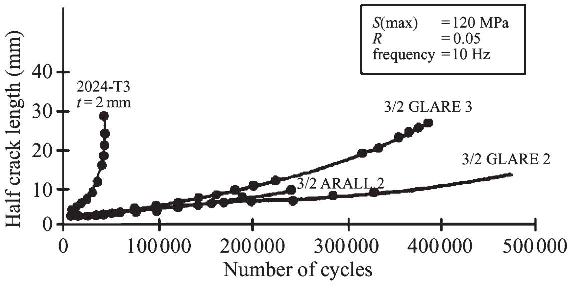


Figure 3
Crack length vs. number of cycles for GLARE and aluminum.

The longitudinal and circumferential riveted joints in a fuselage structure are the locations where fatigue damage will most probably occur. The residual strength of 2024-T3 and GLARE riveted lap joints after fatigue is shown in Fig. 4. The GLARE lap joint shows superior behavior over the aluminum lap joint; a high initial residual strength and slow strength reduction are indicated. Cracks initiate in, and remain

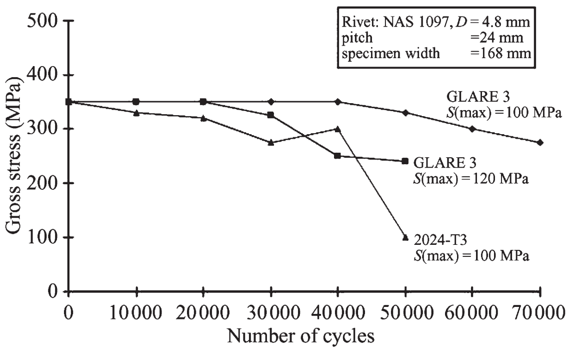


Figure 4
Residual strength of aluminum 2024-T3 and GLARE 3-3/2-0.3 riveted lap joints after fatigue.

limited to, one aluminum layer at the mating surface for a long portion of the life of the material. The 2024-T3 riveted lap joint shows a fast strength decrease once fatigue cracks have initiated. Relatively short inspection intervals are required for aluminum to prevent a situation in which unstable crack extension and reduced residual strength can occur. The results for the GLARE joints indicate that fracture of a riveted lap joint as a result of fatigue damage will never occur under realistic conditions. Damage tolerance is built into the material instead of having to rely on adequate inspection.

3. Fatigue Mechanism in Fiber–Metal Laminates

The fatigue process of fiber–metal laminates can be divided in three phases:

Crack initiation. Because of the lower stiffness of the fiber layers, the aluminum layers in fiber–metal laminates are more heavily loaded, and consequently the crack initiation period in GLARE is usually shorter than in monolithic aluminum. Since an internal stress system is present in the laminate, the initiation period also depends on the operating temperature. When the crack originates from a notch (e.g., hole or saw cut) the initial crack growth rate during the first 1–3 mm of the crack length measured from the notch edge is relatively high. This is related to the stress concentration induced by the notch, but also some crack length and crack opening is needed before the fibers become effective in slowing down the crack growth rate.

Crack growth. During the crack growth period the crack growth rate is almost constant and very low, while the crack growth rate in monolithic aluminum will increase rapidly. Whereas for monolithic aluminum the total fatigue life mainly consists of the crack initiation life, the life for fiber–metal laminates mainly consists of the crack propagation period. Because general bending is present in the structure, e.g., in riveted lap joints, the crack will initiate and grow in one layer only. This fatigue crack remains in the outer layer for a very long time before the crack initiates in the layer underneath. The fibers and intact aluminum layers transfer a significant part of the load over the surface crack and restrain crack opening. If the internal stress system is reversed by a stretching operation after cure or by high peak loads, the crack growth rate will decrease with increasing crack length and crack growth is eventually stopped. High peak loads during variable amplitude fatigue loading are therefore more effective in slowing down crack growth in fiber–metal laminates compared to monolithic aluminum.

Failure. Because the fibers stay intact during fatigue, the residual strength of a cracked GLARE laminate is very high. The material can withstand very large cracks before failure occurs.

As already mentioned, fatigue crack growth in the aluminum layers is accompanied by growth of the delaminated area around the crack due to cyclic shear stresses present in the adhesive which transfer the load from the metal to the fibers (Fig. 2). The magnitude of the bridging stresses in the fibers is related to the crack opening displacement and the length over which the bridging fibers are elongated, i.e., the delamination length. Since the crack opening displacement decreases from a maximum at the crack center to zero at the crack tip, the crack bridging stresses also decrease to zero at the tip. Marissen (1988) introduced a relation for the contour of the crack opening displacement in accordance with a constant fiber bridging stress over the whole crack length. He assumed an elliptical delamination that grows in such a way that the fiber stresses become equal along the crack flanks. This is not unreasonable since a local higher fiber stress will cause locally more delamination which reduces the fiber stress. Although a decrease of the bridging stresses very close to the tips must take place, the fiber stresses will become constant away from the crack tips.

In an optimal situation, the crack opening should be minimal. This would be the case for an infinitely stiff bond between the aluminum and fiber layers. However, in a real situation, due to local shear deformation of the adhesive and delamination, more crack opening will occur. This is schematically represented in Fig. 5. During the fatigue process the adhesive is loaded with cyclic shear stresses which cause the delamination to grow. In most cases this delamination is found to be elliptical in shape. As a consequence, the bridging stress in the fibers will decrease, resulting in lower cyclic shear stresses in the adhesive. This means a reduction of the delamination growth rate. Increase of the crack length and crack opening will result in higher bridging stresses and shear stresses. This means that the crack growth in fiber–metal laminates is governed by a balance

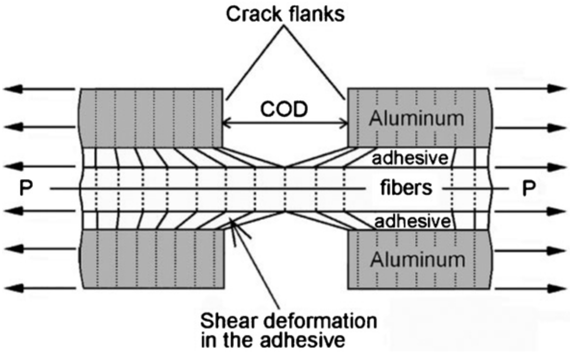


Figure 5
Shear deformation in the adhesive layer.

between the delamination growth rate on the one hand and the bridging stress on the other.

4. *Applications and Future Development*

The first flying GLARE part was a bonded patch repair installed on a C-5A Galaxy aircraft by the US Air Force in October 1995. The first flying primary GLARE structure was a fuselage demonstrator panel on an Airbus A310 in 1999. The first civil series production applications were impact-critical cargo floors of Boeing and Airbus aircraft, the bulkhead of the Bombardier Learjet, and a blast-resistant aircraft container. New applications outside the aerospace industry will be explored and new variants will be added to the fiber–metal laminates family, e.g., high-temperature laminates with titanium and carbon fibers.

Bibliography

Chen J L, Sun C T 1989 Modeling of orthotropic elastic–plastic properties of ARALL laminates. *Compos. Sci. Technol.* **36**, 321–37

Jongebreur A A, West J, Homan J J, van Veggel L H 1992 Structural aspects of ARALL 3 compared with ARALL 2 wing structures. In: Beukers A, De Jong Th, Sinke J, Vlot A, Vogelesang L B (eds.) *Fatigue of Aircraft Materials*. Delft University Press, The Netherlands

Krishnakumar S 1994 Fiber metal laminates—the synthesis of metals and composites. *Mater. Manufact. Processes* **9**, 295–354

Marissen R 1988 Fatigue crack growth in ARALL, a hybrid aluminium–aramid composite material, crack growth mechanisms and quantitative predictions of the crack growth rate. Dissertation, Delft University of Technology, The Netherlands

Roebroeks G 1991 Towards GLARE—the development of a fatigue insensitive and damage tolerant aircraft material. Dissertation, Delft University of Technology, The Netherlands

Vašek A, Polák J, Kozák V 1997 Fatigue crack initiation in fibre–metal laminate GLARE 2. *Mater. Sci. Eng.* **A234–236**, 621–4

Vlot A, Vogelesang L B, de Vries T J 1999 Towards application of fibre metal laminates in large aircraft. *Aircraft Eng. Aerospace Technol.* **71** (6), 558–70

Vogelesang L B, Gunnink J W 1986 ARALL: a materials challenge for the next generation of aircraft. *Mater. Design.* **7** (6), 287–300

Vogelesang L B, Marissen R, Schijve J 1981 *A New Fatigue Resistant Material: Aramid Reinforced Aluminium Laminate (ARALL)*. Report LR-322, Delft University of Technology, The Netherlands

A. Vlot and R. C. Alderliesten
Delft University of Technology, The Netherlands

Fiber Orientation in Composites: Influence of Fluid Flow

A short-fiber composite consists of a polymer, metal, or ceramic matrix reinforced with short, discontinuous fibers. When a short-fiber composite is molded or extruded, the flow gives a pattern of preferential orientation to the fibers. This fiber orientation distribution determines the properties of the composite, since the fibers mainly reinforce the matrix along the axes.

This article discusses techniques for modeling flow-induced fiber orientation. The objective is to predict the orientation distribution in a processed part. This allows one to predict the mechanical performance of the part and, if necessary, alter the process to improve part performance. This type of calculation is already used for injection molded polymer matrix composites, to ensure suitable structural performance and dimensional accuracy before constructing a mold. More extensive reviews are provided by Tucker and Advani (1994) and Papathanasiou and Guell (1997).

1. *Describing Orientation*

The first step in modeling fiber orientation is to describe the orientation of the fibers at any point in a part. The fibers can usually be regarded as rigid cylinders, so the orientation of a single fiber is described by a unit vector **p** pointing along the fiber axis. The set of all possible directions **p** defines the orientation space, which is the surface of a unit sphere.

In fact, a single fiber direction is not an adequate description for most composites. Figure 1 shows a polished section of an injection molded composite. The glass fibers in this material are nearly perfect cylinders, so they appear on the cross section as dark ellipses. The near-circular ellipses are fibers whose axes are almost perpendicular to the section plane, and the highly elongated ellipses are fibers whose axes lie nearly parallel to the plane. Incomplete ellipses occur when the section plane passes through the end of a fiber.

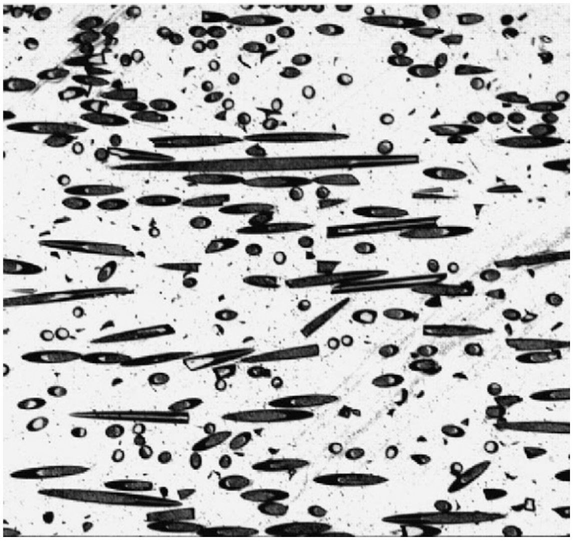


Figure 1
Micrograph of a polished cross section of an injection-molded short-fiber composite. The glass fibers are approximately 10 μm in diameter.

Figure 1 demonstrates that the fibers in any small region of a short-fiber composite may have very different orientations from one another, so it is necessary to describe the *distribution* of fiber orientations. A very complete description is provided by the probability density function $\psi(\mathbf{p})$, defined such that $\psi(\mathbf{p})\mathrm{d}\mathbf{p}$ is the probability that a fiber is oriented within the solid angle $\mathrm{d}\mathbf{p}$ centered on the direction $\mathbf{p}$. ψ is defined on the orientation plane, i.e., on the surface of a unit sphere. Real composite parts also have significant spatial variations in fiber orientation, so ψ is typically taken to be a function of orientation $\mathbf{p}$, position $\mathbf{x}$, and time t.

The amount of information needed to describe $\psi(\mathbf{p}, \mathbf{x}, t)$ for a realistic part makes the calculation of ψ prohibitively expensive. Instead, the standard approach is to calculate moments of the distribution function. These moments are tensorial quantities, and are called the orientation tensors, configuration tensors, or moment tensors (Advani and Tucker 1987). The second-order orientation tensor is the average of the tensor product $\mathbf{pp}$ over all directions $\mathbf{p}$, weighted by the probability density function ψ:

$$A = \int \mathbf{pp}\psi(\mathbf{p})\mathrm{d}\mathbf{p} \tag{1}$$

One may also think of A as the average value of $\mathbf{pp}$ for all the fibers in a small region. The discussion will also require the fourth-order tensor, which is:

$$A = \int \mathbf{pppp}\psi(\mathbf{p})\mathrm{d}\mathbf{p} \tag{2}$$

The orientation tensors reveal the main directions of orientation (through the eigenvectors of $\mathbf{A}$) and show how aligned the fibers are about these directions. For example, if all of the fibers are aligned in the x_1 direction, then the second-order orientation tensor is:

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{3}$$

while for a three-dimensional random orientation, the orientation tensor is

$$\mathbf{A} = \begin{bmatrix} 1/3 & 0 & 0 \\ 0 & 1/3 & 0 \\ 0 & 0 & 1/3 \end{bmatrix} \tag{4}$$

The off-diagonal components of $\mathbf{A}$ are nonzero when the principal axes of orientation are nonaligned with the coordination axes. $\mathbf{A}$ is a symmetric tensor and $A_{11} + A_{22} + A_{33} = 1$, so $\mathbf{A}$ has only five independent components. One must still solve for the five independent components of $\mathbf{A}$ as functions of $\mathbf{x}$ and t, but this is far less information than the probability density function ψ.

2. *Orientation Dynamics*

The basic solution for the rotational motion of a fiber in a deforming fluid was derived by Jeffery (1922). He considered a single, rigid ellipsoidal particle in a Newtonian fluid, with negligible inertia and body forces. Subsequent theoretical and experimental work showed that any rigid, axisymmetric particle will follow Jeffery's equation, provided the correct value of the axis ratio was used. Jeffery's result shows that the centroid of the fiber moves with the unperturbed fluid velocity, and the fiber axis $\mathbf{p}$ reorients according to

$$\dot{\mathbf{p}} = \mathbf{W} \cdot \mathbf{p} - \lambda(\mathbf{D} \cdot \mathbf{p} - \mathbf{D}{:}\mathbf{ppp}) \tag{5}$$

Here, the dot denotes a time derivative following the fiber, $D_{ij} = (\delta v_i/\delta x_j + \delta v_j/\delta x_i)/2$ is the rate-of-deformation tensor, $W_{ij} = (\delta v_i/\delta x_j - \delta v_j/\delta x_i)/2$ is the vorticity tensor, and $\lambda = (r_e^2 - 1)/(r_e^2 + 1)$ is a constant that depends on the fiber shape. r_e is called the equivalent axis ratio; for a slender fiber it is approximately equal to the fiber length divided by the fiber diameter.

Fibers in very dilute suspensions follow Jeffery's equation exactly. Shearing flows align fibers in the flow direction, while stretching deformations align fibers in the direction of stretching. This latter rule means that fibers in a converging flow rotate towards the flow direction, while fibers in a diverging flow rotate to lie across the flow.

Practical composites always involve high concentrations of fibers, where fiber–fiber interactions come into play. Theoretical modeling of these interactions is quite difficult, and most computations use a phenomenological model proposed by Folgar and Tucker (1984). The effect of interactions is modeled as a rotary diffusion process (assuming that fibers bump into one another in a nearly random way), and the rotary diffusivity is assumed to be proportional to the scalar shear rate $\dot{\gamma} = (2\mathbf{D}{:}\mathbf{D})^{1/2}$. Adding this term to Jeffery's equation and combining the result with Eqn. (1) gives an evolution equation for the second-order orientation tensor:

$$\dot{\mathbf{A}} = (\mathbf{A}\cdot\mathbf{W} - \mathbf{W}\cdot\mathbf{A}) + \lambda(\mathbf{A}\cdot\mathbf{D} + \mathbf{D}\cdot\mathbf{A} - \mathbf{D} : \mathbb{A}) + 2C_I\dot{\gamma}(\mathbf{I} - 3\mathbf{A}) \qquad (6)$$

Here, the dot represents a material derivative and **I** is the unit tensor. C_I is the interaction coefficient, a dimensionless quantity describing the strength of fiber–fiber interactions that is adjusted to match experimental data. Increasing the value of C_I causes the steady-state orientation to be less perfectly aligned. C_I values between 10^{-2} and 10^{-3} are used to match data for fiber-reinforced polymers used in injection molding.

Equation (6) could be solved to follow the evolution of the second-order orientation tensor **A**, except that it contains the fourth-order orientation tensor $\mathbb{A}$. The corresponding evolution equation for $\mathbb{A}$ contains a sixth-order tensor, so some approximation is needed to close the set of equations. The usual strategy is to develop a closure formula that approximates $\mathbb{A}$ in terms of **A**, and substitute it into Eqn. (6). The simplest useful form is the quadratic closure:

$$\mathbb{A} = \mathbf{AA} \qquad (7)$$

however, this is not always very accurate.

A variety of other closure approximations have been developed, the most accurate being the orthotropic family of closures (Cintra and Tucker 1995), and the natural closure (Dupret *et al.* 1997). Once a closure approximation has been chosen and the velocity field is known, Eqn. (6) can be integrated to find the orientation as a function of position and time.

3. *Process Models*

The preceding section showed that the fiber orientation is carried along with the flow of the fiber/liquid mixture, and that it evolves in response to the rate of deformation **D** and the vorticity **W**. Thus, fiber orientation changes are driven by velocity and by velocity gradients. To model the orientation in a process like extrusion or molding, one must solve the flow problem to find the velocity field and then use those velocities in Eqn. (6) to solve for orientation.

Inertia, body forces, and density change are typically negligible in the process of interest, so the governing mass and momentum balance equations for the flow problem are

$$\nabla\cdot\mathbf{v} = 0 \qquad (8)$$

$$-\nabla p + \nabla\cdot\tau = 0 \qquad (9)$$

Here p is the fluid pressure and τ is the extra stress tensor. Many processing problems also involve significant temperature changes, temperature-dependent viscocity, and solidification, in which case the energy equation is also needed to provide the temperature field. Solving Eqns. (6), (8), and (9), subject to suitable initial and boundary conditions, provides the desired results. These solutions are usually found numerically.

Next, several example solutions will be shown. All of these calculations were performed by VerWeyst (1998), using a finite element code that uses pressure, velocity, and the components of the orientation tensor **A** as nodal degrees of freedom. The orientation calculations are all time-dependent and start from an initial condition of random orientation, but eventually the orientation solution reaches a steady state. The fiber orientation results are displayed here by drawing a line at each node, in the direction of the largest eigenvector of **A**. The direction of this line is the principal direction of orientation, and the length of the line indicates the amount of fiber alignment around that direction: a long line corresponds to near-perfect alignment, while a very short line corresponds to nearly random fibers.

Figure 2 shows a steady-state solution for flow and fiber orientation in an axisymmetric contraction, where a tube of diameter 4.5 joins a tube of diameter 1. This is a prototype for flow in extrusion dies and capillary rheometers. The streamlines are for an isothermal Newtonian fluid with zero Reynolds number, and the computational mesh extends well upstream and downstream of the region shown in the figure. The converging streamlines at the contraction correspond to strong stretching in the flow direction, and the fibers become strongly aligned by this stretching flow as they approach the smaller tube. There is a slow recirculation in the corner, not apparent in the streamline plot, but easily seen in the fiber orientation pattern. The overall result of this flow is to produce a high degree of flow-direction alignment downstream of the contraction.

A second example concerns radial flow in a thin disk, entering from a tube connected to the center of the disk. Figure 3 shows the overall geometry, and Fig. 4 shows the streamlines and steady-state orientation pattern near the center of the disk. This flow is a prototype for injection molding, particularly for flow near a gate. The streamlines again correspond to an isothermal Newtonian fluid. The orientation pattern in this case is quite different from Fig. 2. Now, in the vertical tube, shearing has aligned the fibers in the flow direction. However, in the disk, the streamlines

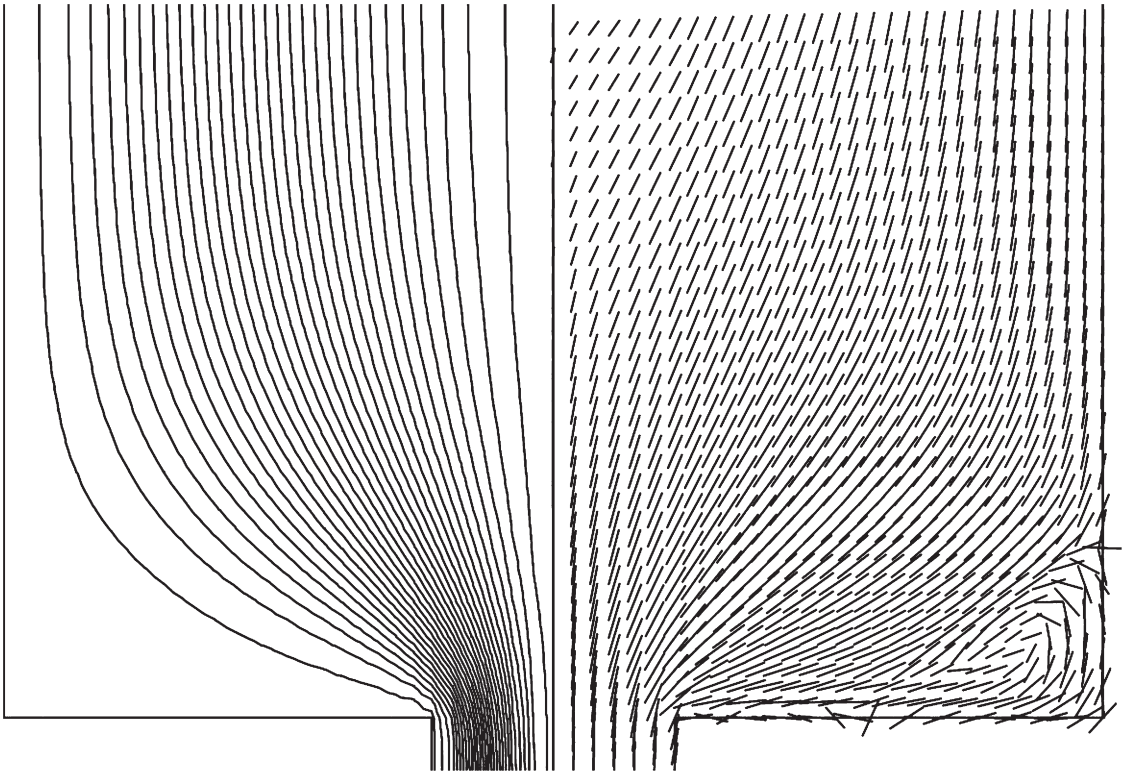


Figure 2
Computed streamlines and fiber orientation pattern for an axisymmetric contraction. Flow is from top to bottom.

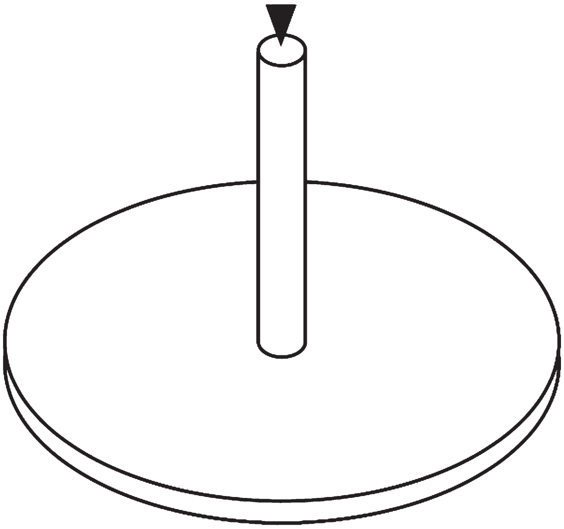


Figure 3
Geometry for the flow in a center-gated disk.

are diverging, which provides a strong stretching deformation in the hoop direction.

Figure 4 shows that, once the fibers enter the disk, they quickly rotate so that they are perpendicular to the flow. (Orientations perpendicular to the flow appear as dots in Fig. 4.) Near the upper and lower walls of the disk, the stretching effect is outweighed by shearing parallel to the walls (due to the no-slip boundary condition), and the fibers there begin to align in the flow direction. Further downstream of the area shown in the figure, an inner core of fibers transverse to the flow is sandwiched between outer shells of flow-aligned fibers. This shell/core structure is typical of injection-molded parts.

A more complex example is shown in Fig. 5 (VerWeyst *et al.* 1999). This is a short rib in a large plaque, 1.5 mm thick, that is injection molded. The rib is long in the direction perpendicular to the page, which allows this problem to be analyzed as a two-dimensional flow. This particular analysis includes all of the factors needed to accurately model injection molding: the polymer viscosity depends on temperature and shear rate, the energy equation is solved at

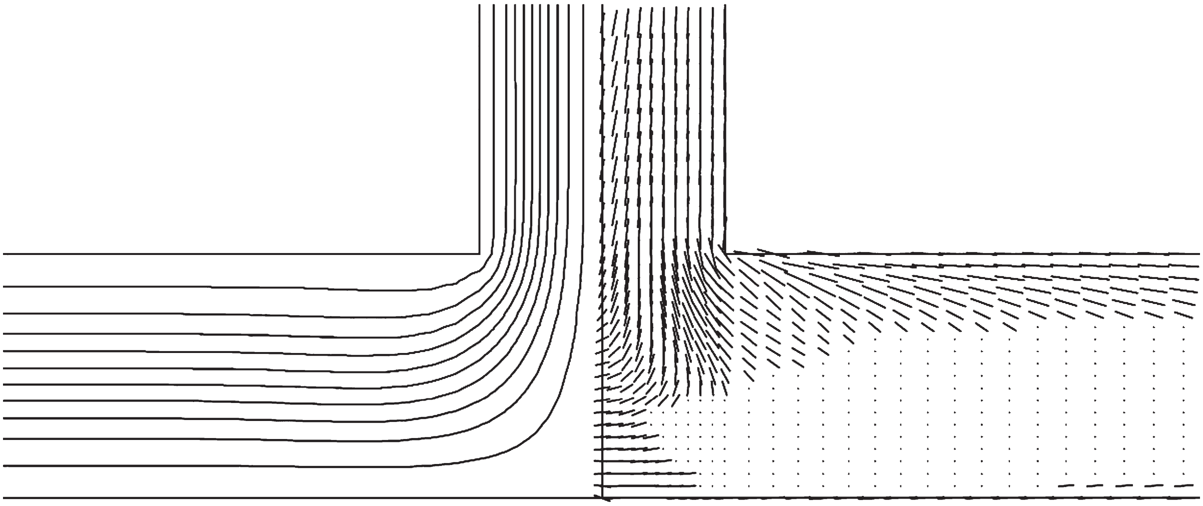


Figure 4
Computed streamlines and fiber orientation pattern for flow in a center-gated disk. Flow enters from the top and exits at the sides.

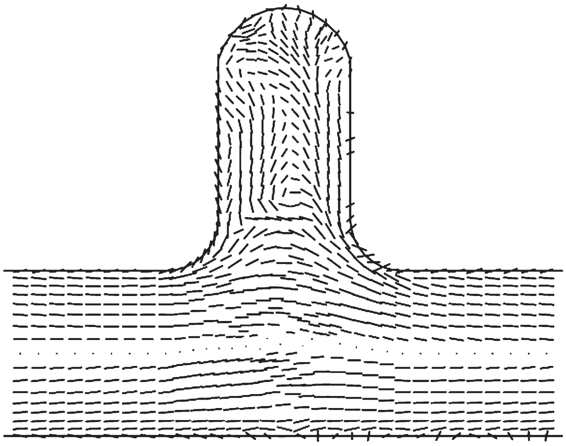


Figure 5
Computed fiber orientation pattern for a cross-flow rib in an injection molded plaque. The rib is filled from the left, and flow exits at the right.

each time step to provide temperatures, and the analysis starts with an empty mesh and includes the moving flow front as the polymer/fiber mixture fills the mold cavity.

The orientation results in the plaque show the core/shell structure typical of injection molding, although in this case the fibers at the midplane are randomly oriented. Within the rib, the orientation pattern is more complex. Figure 5 shows a scanned image of a cross section from a molded part with this same geometry. The dark spots in the figure are glass fibers and the white background is the polycarbonate matrix. The fibers in this image have the same orientation pattern as predicted by the calculations, a fact confirmed by more quantitative comparisons (VerWeyst *et al.* 1999).

The preceding examples solve the flow problem using the full two- or three-dimensional mass and momentum balance equations. However, in some processes, a simpler set of equations can be used to solve for the flow field. For polymer injection molding, one solves the generalized Hele–Shaw formulation (Hieber and Shen 1980). Gupta and Wang (1993) and Crochet *et al.* (1994) discuss how this type of flow solution is coupled with the fiber orientation equations.

In compression molding, the fibers are much longer than the part thickness, so the fibers are constrained to lie parallel to the mold surface. Then the orientation can be modeled as two-dimensional, and can be combined with a specialized solution of the flow problem. Advani and Tucker (1990) and Osswald and Tseng (1994) discuss models for compression molding.

4. *Additional Issues*

A number of other modeling issues must be addressed to make the simulation of fiber orientation useful. First, it is essential to be able to translate the orientation results into mechanical properties of the finished part. Consider an injection molded part, and suppose that a finite element mesh has been used to simulate mold filling and to predict fiber orientation. A micromechanics model is then applied to each element to predict its stiffness and thermal expansion, after which one may perform a finite element structural analysis of the part. Fortunately, the orientation tensor provides exactly the information about fiber orientation that is needed to predict mechanical properties (Advani and Tucker 1987). Gupta and Wang (1993) and Zheng *et al.* (1999) give examples of this type of calculation.

A second issue is that changes in the fiber orientation alter the rheological properties of the fiber suspension. The examples in the preceding section ignored this phenomenon, calculating the flow field independently of the fiber orientation, and using that flow field to drive the orientation calculation. In fact, the presence of the fibers may alter the flow field, a phenomenon demonstrated experimentally by Libscomb *et al.* (1988). These same workers showed that the fiber suspension flow could be accurately modeled by using a constitutive equation for the extra stress τ that has the form:

$$\tau = 2\eta \mathbf{D} + 2\eta N_p A : \mathbf{D} \tag{10}$$

Here, η is the viscosity of the fluid matrix, and N_p is a dimensionless quantity that depends on the fiber volume fraction and aspect ratio. Theories for predicting N_p, as well as other aspects of fiber suspension rheology, are reviewed by Petrie (1999).

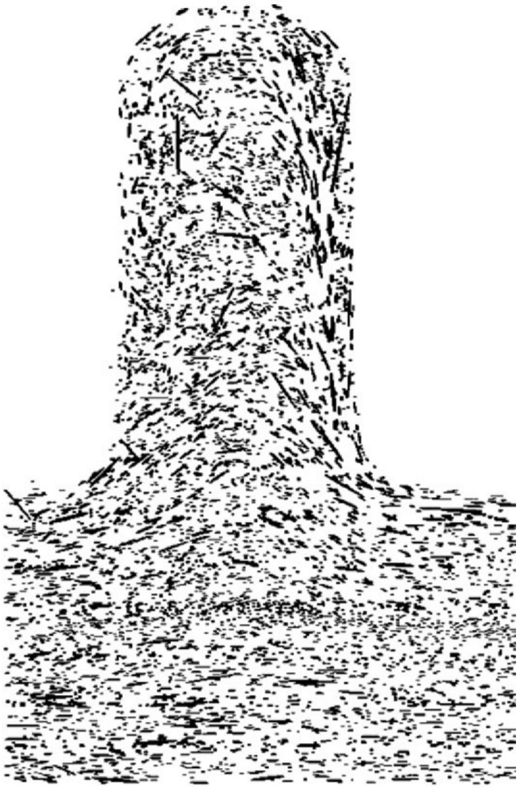


Figure 6
Polished cross section of a cross-flow rib in an injection molded plaque. Compare to Fig. 5.

Figure 7 shows the velocity field and orientation pattern in an axisymmetric contraction for $N_p = 6$. This is a fully coupled equation: Eqn. (10) is used to compute stresses in the suspension, and the velocity field and orientation field are solved simultaneously. Comparing this to Fig. 2, which is the same geometry, shows that the presence of the fibers greatly enlarges and strengthens the corner vortex. The fibers resist stretching in the direction of their axes, but provide little resistance to shearing parallel to their axes. This causes the angle of the entrance streamlines to narrow, so that the elongational deformation is spread out over a greater length of the large tube. While the coupling between flow and orientation is important in this contraction flow, it is a weak effect for flows in narrow gaps (typical of injection and compression molding), and may be neglected in many problems of practical interest.

A third issue is the breakage of fibers during flow. Generally speaking, large stresses during flow lead to fiber breakage and reduction of the average fiber length, with attendant reduction in the mechanical properties of the composite. Some simple models provide a general understanding of the mechanisms of fiber breakage (Papathanasiou and Guell 1997). However, current practice is to design molds, runners, and gates to minimize fiber breakage, based on empirical rules, rather than to attempt a detailed prediction of breakage.

5. *Other Modeling Approaches*

This discussion has focused on the use of the second-order orientation tensor **A** as the primary variable for describing and calculating fiber orientation, an approach that has proved quite effective, and is incorporated in commercial software for modeling polymer injection molding. However, some other modeling approaches have been developed recently and deserve mention.

In the *Brownian configuration field* method (Fan *et al.* 1999), one calculates a field of orientation directions $\mathbf{p}(\mathbf{x}, t)$. The governing equation is Jeffery's equation, Eqn. (5), with an additional term representing the stochastic force from fiber–fiber interactions. The same stochastic force is applied to every fiber in the field. This method uses many fields of orientation, each with a different initial condition and a different stochastic force. By choosing the stochastic forces properly, the average result becomes equivalent to the rotary diffusion model proposed by Folgar and Tucker. Once the results have been calculated, the fiber directions from all fields at a given node can be combined to approximate the orientation distribution function ψ, or they can be averaged to compute the orientation tensor **A**. The

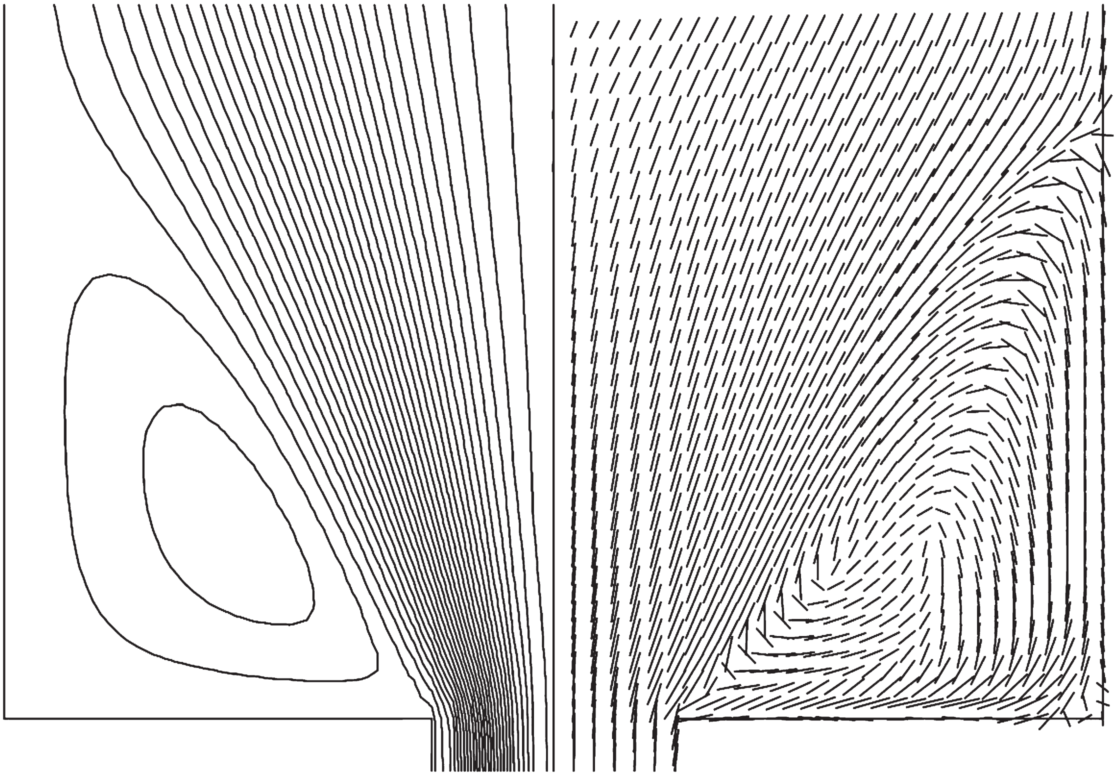


Figure 7
Computed streamlines and fiber orientation pattern for an axisymmetric contraction with fully coupled flow and orientation ($N_p = 6$). Flow is from top to bottom. Compare to Fig. 2, which corresponds to $N_p = 0$.

Brownian configuration field calculation is more expensive than calculating **A** directly, but it does not require a closure approximation so it is potentially more accurate.

A second, very different, approach is the *direct simulation* of a suspension with multiple fibers (Fan *et al.* 1999). Rather than using Jeffery's equation, the Navier–Stokes equation is solved at the microscopic level, in the spaces between the fibers. Each fiber is treated as a rigid cylinder that has translational and rotational motion. At each time step, the method solves for the velocity distribution between the fibers, as well as the rates of translation and rotation that cause each fiber to experience no net force or moment. This is an extremely challenging calculation, and direct simulations are presently limited to about 100 fibers in a box with periodic boundary conditions.

Thus, direct simulation is not useful for modeling entire extrusion or molding processes. The advantage of direct simulation is that it requires no approximations about fiber–fiber interactions. Instead, the effects of interactions arise naturally, as the fibers bump into one another or experience long-range hydrodynamic forces. Direct simulation provides a way to make theoretical predictions of the interaction coefficient C_I, and recent calculations show promising agreement with experiments.

6. Summary

Flow-induced fiber orientation, generated during processing, determines the mechanical and physical behavior of short fiber composites. Fiber orientation distributions can be predicted by solving for the velocity field during flow, and by using a model of orientation dynamics. Current practice is to use the second-order orientation tensor (or moment tensor) **A** as the primary variable to represent local fiber orientation. This quantity captures both the principal directions of orientation and the extent of alignment about those directions. Governing equations for the evolution of **A** can be developed from the basic solution for single-fiber motion, and the resulting equations can be solved numerically for complex flow fields. The results of the orientation model can then be used to predict the mechanical properties and structural performance of the finished part.

See also: Composite Materials: Overview#; Polymer Injection Molding: Modeling for Properties#; Liquid Crystalline Polymers: An Introduction#

Bibliography

Advani S G, Tucker C L III 1987 The use of tensors to describe and predict fiber orientation in short-fiber composites. *J. Rheol.* **31**, 751–84

Advani S G, Tucker C L III 1990 A numerical simulation of short-fiber orientation in compression molding. *Polym. Compos.* **11**, 164–73

Cintra J S Jr, Tucker C L III 1995 Orthotropic closure approximations for flow-induced fiber orientation. *J. Rheol.* **39**, 1095–122

Crochet M J, Dupret F, Verleye V 1994 Injection molding. In: Advani S G (ed.) *Flow and Rheology in Polymer Composites Manufacturing*. Elsevier, Amsterdam, pp. 415–63

Dupret F, Verleye V, Languillier B 1997 Numerical prediction of the molding of composite parts. In: Advani S G, Siginer D A (eds.) *Rheology and Fluid Mechanics of Nonlinear Materials*. FED-Vol. 243/MD-Vol. 78. ASME, New York, pp. 79–90

Fan X, Phan-Thien N, Zheng R 1999 Simulation of fibre suspension flows by the Brownian configuration field method. *J. Non-Newtonian Fluid Mech.* **84**, 257–74

Folgar F, Tucker C L III 1984 Orientation behavior of fibers in concentrated suspensions. *J. Reinf. Plastic Compos.* **3**, 98–119

Gupta M, Wang K K 1993 Fiber orientation and mechanical properties of short-fiber-reinforced injection molded composites: Simulated and experimental results. *Polym. Compos.* **14**, 367–82

Hieber C A, Shen S F 1980 A finite-element/finite-difference simulation of the injection molding filling process. *J. Non-Newtonian Fluid Mech.* **7**, 1–32

Jeffery G B 1922 The motion of ellipsoidal particles immersed in a viscous fluid. *Proc. R. Soc.* **A102**, 161–79

Lipscomb G G, Denn M M, Dur D U, Boger D V 1988 The flow of fiber suspensions in complex geometries. *J. Non-Newtonian Fluid Mech.* **26**, 297–325

Osswald T A, Tseng S-C 1994 Compression molding. In: Advani S G (ed.) *Flow and Rheology in Polymer Composites Manufacturing*. Elsevier, Amsterdam, pp. 361–413

Papathanasiou T D, Guell D C (eds.) 1997 *Flow-induced Alignment in Composite Materials*. Woodhead, Cambridge

Petrie C J S 1999 The rheology of fibre suspensions. *J. Non-Newtonian Fluid Mech.* **87**, 369–402

Tucker C L III, Advani S G 1994 Processing of short-fiber systems. In: Advani S G (ed.) *Flow and Rheology in Polymer Composites Manufacturing*. Elsevier, Amsterdam, pp. 147–202

VerWeyst B E 1998 *Numerical Predictions of Flow-induced Fiber Orientation in 3D Geometries*. PhD Thesis. University of Illinois at Urbana-Champaign, USA

VerWeyst B E, Tucker C L III, Foss P H, O'Gara F J 1999 Fiber orientation of 3D injection moulded features: Prediction and experiment. *Int. Polym. Proc.* **14**, 409–20

Zheng R, Kennedy P, Phan-Thien N, Fan X-J 1999 Thermoviscoelastic simulation of thermally and pressure-induced stresses in injection moulding for the prediction of shrinkage and warpage for fibre-reinforced thermoplastics. *J. Non-Newtonian Fluid Mech.* **84**, 159–90

C. L. Tucker III
University of Illinois, Urbana, Illinois, USA

Fiber Strength

Tensile strength is one of the most important physical characteristics of reinforcement fibers. Accurately quantifying strength and variability in strength is therefore of great importance. Fiber strength measurements are used by fiber and composite developers and manufacturers as a means of predicting composite properties and to assess the consistency of different lots of fibers.

Very high strength has been achieved by inorganic materials in fiber form. For instance, a tensile strength of >1 GPa in monolithic ceramics is considered exceptional. By contrast, multifilament tows of SiC and Al_2O_3 are commercially available with strengths of 3–4 GPa. Commercial PAN-based carbon fibers exhibit even higher strengths of 5–7 GPa. Large-diameter monofilaments such as SiC fibers prepared by chemical vapor deposition (CVD) and single-crystal sapphire fibers also have strengths in excess of 5 GPa.

Two types of methods are used to measure fiber strength: single filament testing and dry bundle or tow testing. Single filament testing is the most direct route to determining fiber strength, and thereby provides the best data for understanding the fundamentals of behavior and variability. On the negative side, this technique is somewhat difficult because handling of small-diameter, high-modulus fibers and measuring small loads and small filament diameters requires special care to avoid errors that can compromise the accuracy of the test. Testing of tows is often used so that these issues can be avoided. Tow testing is attractive in that it allows the evaluation of hundreds or thousands of filaments simultaneously.

The variability of strength within a sample of fibers is very important. Models indicate that strength of composites is determined not only by the mean tensile strength of the reinforcing fiber, but also by the distribution of strength of fibers. This occurs since the failure of a few weak fibers within the composite at low loads can lead to failure of the entire composite. Weibull statistics are commonly used to predict strength at different tested volumes (gauge length, area), to predict the strength of bundles of fibers, and, ultimately, to predict the strength of fiber-reinforced composites.

1. Single-filament Strength Testing: Procedural Issues

Fibers used for the reinforcement of polymer, metal, and ceramic composites behave as brittle, elastic solids, i.e., fibers exhibit a linear stress–strain curve. Thus, strength is defined as the load at failure divided by the cross-sectional area of the fiber. Single-filament testing is performed as follows. A single filament is selected from the bundle of filaments, carefully aligned and mounted on a special tab (typically of thin cardboard) with an open slot in the center corresponding to the gauge length, typically 25 mm. The filament is mounted on the tab with a suitable adhesive strong enough to prevent the fiber from pulling out during testing but not so stiff that fibers break at the edges of the adhesive. It is critical to mount high-modulus fibers with as perfect alignment as possible to prevent bending stresses which cause premature failure at the edge of the adhesive. The ends of the tab are gripped in a testing machine, the sides of the tab are carefully cut or burned away, and the fiber is tested to failure at a constant strain rate, usually 0.002–0.02 min^{-1}. For the testing of large-diameter monofilaments, the filament may be directly loaded into the grips of the test machine. The grips must be firm enough to grip the fiber and prevent it from pulling out during testing, but soft enough not to damage the fiber or create stress concentrations which cause premature failure in the grips. Soft aluminum foil is one suitable grip material. Direct gripping using rubber-faced grips has also been used for testing of small-diameter filaments (Wilson 1997).

The second part of single-filament testing is determining the cross-sectional area of the fiber. The measurement of fiber area, although conceptually simple, is not straightforward. A round-robin test undertaken by a joint government–industry task group has identified a number of shortcomings and concerns with the standard test method ASTM 3379-75 (Hurst *et al.* 1998). This study found that the measurement of filament diameter is the major source of error in fiber strength determinations.

The small size of fibers makes an accurate determination of fiber cross-section difficult. Most multifilament tow fibers are 7–15 μm in size; thus, a very small error in diameter measurement can have a large effect on calculated fiber strength. For instance, for a 12 μm diameter fiber, a measurement error of only 0.6 μm in diameter will change the calculated strength by 10%. If calibrated, scanning electron microscopy (SEM) can produce very accurate diameter measurements, but SEM is too time-consuming and costly for routine testing. Optical microscopes, often in combination with video image analysis systems, are frequently used for diameter measurement but are limited by their ultimate resolution of ~0.5 μm. Laser diffraction patterns from opaque fibers can also be used for diameter measurement. The measurement of monofilaments is more straightforward because of their large diameter. Optical microscopy or mechanical micrometers can be used to measure the diameter of large-diameter monofilaments.

Additional complications in diameter measurement also exist. During testing, high-modulus fibers almost always shatter due to the release of stored strain energy at fracture, causing the original fracture location to be lost. The diameter of many, if not most, fibers varies slightly down their length, so this effect can make it impossible to determine the actual stress at the point of fracture. A further complication is that fibers may not be round but oval or irregular in cross-section. In this case, it is critical to measure fiber cross-section end-on rather than longitudinally.

In many cases, for instance to assess lot-to-lot variability or to measure degradation of properties after exposure to certain process conditions, it is appropriate to use average strength. ASTM 3379-75 calls for the measurement of a "representative" number of filament diameters from planimetric sample cross-sections of composites to determine a mean filament diameter. The mean diameter is used to calculate the individual fiber breaking stresses. The use of mean diameter is sufficient for the determination of *average* fiber strength; however, the use of mean diameter rather than individual fiber diameters means that individual fiber strength values will not be strictly accurate. This effect must be taken into account during statistical analysis of the strength data.

2. Statistics of Fiber Fracture: Weibull Theory

As brittle materials, the strength of fibers is controlled by the strength of the "weakest link," i.e., fracture occurs at the largest flaw present in the sample. This type of behavior is described by Weibull statistics, in which flaws are assumed to be randomly distributed throughout the tested volume. The application of Weibull theory to high-modulus fibers has been reviewed by van der Zwaag (1989). The probability of failure of a material is given by:

$$P = 1 - \exp[-V/V_0(\sigma/\sigma_0)^m] \qquad (1)$$

where m is the Weibull modulus, V is the tested volume, σ is the failure strength, and V_0 and σ_0 are scaling constants.

Reinforcing fibers typically have a Weibull modulus between 4 and 15. The higher the Weibull modulus, the narrower the distribution of fiber strengths within a sample. A convenient method of estimating the Weibull modulus is the equation $m = 1.2/\text{CV}$, where CV is the coefficient of variation of strength at a single gauge length. By rearranging and taking the natural logarithm of both sides of Eqn. (1), the following equation is obtained:

$$\ln[\ln(1/1-P)] - \ln V/V_0 = m \ln \sigma - m \ln \sigma_0 \qquad (2)$$

For a constant tested volume (equivalent to constant gauge length for fibers), Eqn. (2) is reduced to:

$$\ln\,[\ln\,(1/1-P)] = m \ln \sigma + k \qquad (3)$$

where k is a constant. The Weibull modulus, m, is then determined graphically as the slope of the "Weibull plot" of $\ln[\ln(1/1-P)]$ against $\ln\sigma$. For a set

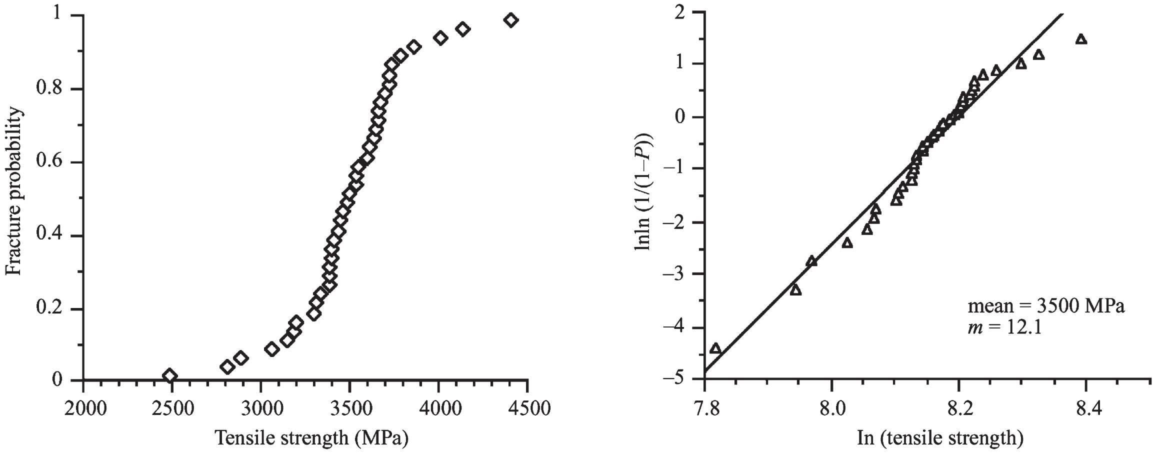


Figure 1
Fracture probability and Weibull plot of Nextel 610 fiber. The fracture probability has sinusoidal form. The Weibull plot is linear with a slope equal to the Weibull modulus ($m = 12$).

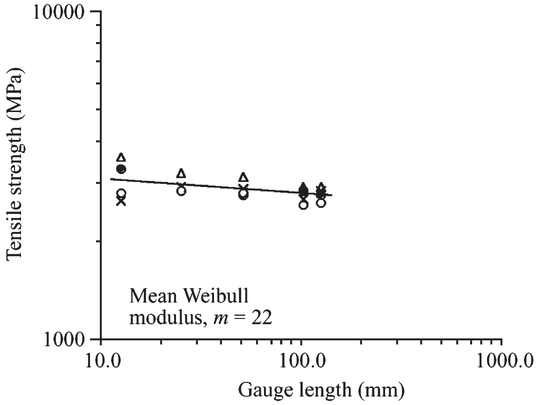


Figure 2
Tensile strength of four lots of Nextel 610fiber at five different gauge lengths. The line is the least squares fit of the mean strength of all four lots ($m = 22$).

of fracture data, the samples are ranked from weakest to strongest, and the probability of fracture, P, is most accurately estimated as $P_i = (i–0.5)/n$, where P_i is the rank of the ith specimen and n is the number of specimens tested. If the statistics of fiber fracture conform to Weibull statistics, the strength data forms a straight line. Figure 1 shows the distribution of single-filament strength for Nextel 610 Ceramic Oxide Fiber, an alumina fiber made by 3M (Wilson 1997). On the strength plot, the data have a sinusoidal form. The Weibull plot is shown in Fig. 2: as can be seen, the data conform to Eqn. (2). The Weibull modulus for this fiber is 12.

It is important to test a sufficient number of filaments to obtain an accurate Weibull modulus. One advantage of single-filament testing is that the number of samples available for testing is very large; however, fiber testing can be very time-consuming, so a tendency exists to minimize the number of fibers tested. Calculations from randomly generated data sets indicate that for a set size of 10 tests, the standard deviation in calculated Weibull modulus may be 30% or more (van der Zwaag 1989). It is recommended that a minimum of 20, and preferably 30 or more, filaments be tested to ensure an accurate statistical analysis of fiber strength.

Another method of determining Weibull modulus is to measure fiber strength as a function of gauge length. With increasing gauge length, the chance of finding a long flaw increases, so fiber strength decreases. If fracture probability, P, is held constant by measuring mean strength, then Eqn. (2) reduces to:

$$\ln \sigma = -1/m \ln V + k' \tag{4}$$

or, alternatively,

$$\sigma_1/\sigma_2 = (V_2/V_1)^{1/m} = (L_2/L_1)^{1/m} \tag{5}$$

where the strengths at tested volumes V_1 and V_2 are σ_1 and σ_2, respectively, and k' is a constant. For fibers, it is commonly assumed that the tested volume is proportional to gauge length, so gauge length, L, can be substituted for volume, V. Using this approach, the Weibull modulus, m, can be determined by plotting log of mean strength as a function of log of gauge length. The slope of the log–log gauge length plot is equal to $-1/m$. Figure 2 shows a log–log plot of tensile strength as a function of gauge length for

Nextel 610 fiber. The Weibull modulus as calculated from Fig. 2 is 22.

Work carried out at the end of the 1990s has illuminated some issues concerning the classic Weibull analysis given in Eqns. (2) and (4) as applied to fibers. The Weibull modulus of fibers as measured by the strength distribution technique (Eqn. (2)) is usually lower than with the gauge length technique (Eqn. (4)) (Wilson 1997, Lavaste *et al.* 1995). This effect can be observed in Figs. 1 and 2. The reason for the discrepancy is not yet clear. One potential cause is incorrect diameter measurement, either due to the use of mean filament diameter rather than individual diameter or from measurement error. It can be shown that either effect will broaden the distribution of strength and therefore reduce the Weibull modulus. Another issue is the effect of natural variations in fiber diameter on tested volume. Since, for given gauge length, larger diameter fibers have larger tested volumes than smaller fibers, Eqn. (5) is not strictly accurate. Since large-diameter fibers would be expected to have lower strength than smaller fibers, this will also broaden the distribution of strengths at a single gauge length. Lastly, the assumption that flaws are distributed randomly with volume is most likely not accurate. For instance, as discussed below, fibers commonly fracture due to surface flaws rather than flaws distributed through the volume of the fiber. Other issues related to processing (e.g., surface damage) can also cause flaws to be nonrandomly distributed. To consider the effects of fiber diameter on fracture probability, Eqn. (1) can be expanded to:

$$P = 1 - \exp[-(\pi/4)LD^{\gamma}/V_0(\sigma/\sigma_0)^m] \qquad (6)$$

where γ is a scaling constant and D is the fiber diameter. γ should in theory be 2, but several investigators have found that a lower value of γ generates a more accurate prediction of fiber strength (Wagner 1989, Zhu *et al.* 1997).

3. *Strength of Fiber Bundles*

The strength of an ideal bundle of fibers will be less than the mean strength of individual filaments. This occurs because the fibers in the bundle have a distribution of strength; therefore, some fibers will fail at a stress below the mean. The remaining stress will be transferred to the other fibers, increasing their load and reducing the overall load required for failure. The ultimate strength of a bundle of filaments which follow Weibull fracture statistics and have mean strength σ_0 was determined by Coleman (1958):

$$\sigma_u = \sigma_0/[m^{1/m}\exp(1/m)] \qquad (7)$$

where m is the Weibull modulus. As the coefficient of variation in fiber strength increases (i.e., m decreases), the strength of the bundle decreases, as shown in Fig. 3. For instance, for $m=6$ the ideal bundle strength is 62% of the mean fiber strength.

The Weibull modulus of individual fibers in the bundle can ideally be determined from a single bundle test. Chi *et al.* (1984) derived an equation to calculate the Weibull modulus from measured parameters on a stress–strain curve of a bundle of fibers:

$$m = [\ln(\varepsilon_{\max}S_0/L_{\max})]^{-1} \qquad (8)$$

where $L_{\max}$ and $\varepsilon_{\max}$ are the maximum load and strain, respectively, and $S_0 = EAN$, the initial slope of the stress–strain curve, where the number of filaments with modulus E is N having total area A.

Experimentally, it is often difficult to measure the "ideal" bundle strength. In an ideal bundle, each filament is continuous, evenly loaded, and exactly parallel to the load axis during testing; additionally, fracture in one filament is assumed not to affect neighboring fibers. These conditions are often not met with in commercial fibers. Loading hundreds or thousands of filaments equally is also not straightforward. Special arrangements are used to ensure uniform tension on fibers in the bundle. A common procedure is to stretch or hang the fiber tow on a frame under tension. The ends of the tow are infiltrated with epoxy or other suitable adhesive and attached to special end tabs for testing. For some fibers, even with this type of care, measured bundle strength is frequently lower than expected from single filament strength. Multifilament tows all have some degree of entanglement, leading to differences in load between filaments. For oxide fibers in which interfilament friction might be expected to be high, fracture of one filament may cause fracture in adjacent fibers. Oils or lubricants applied to tows have been shown to increase measured strength of oxide fibers; this suggests that friction between filaments causes premature fracture.

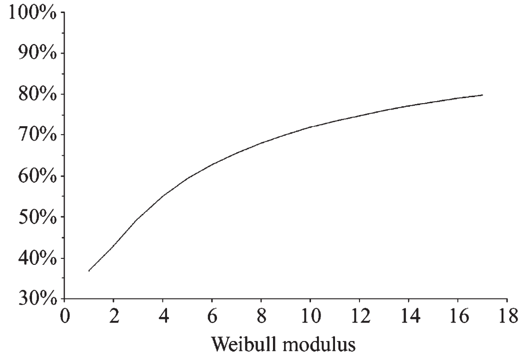


Figure 3
The strength of fiber bundles increases as Weibull modulus increases.

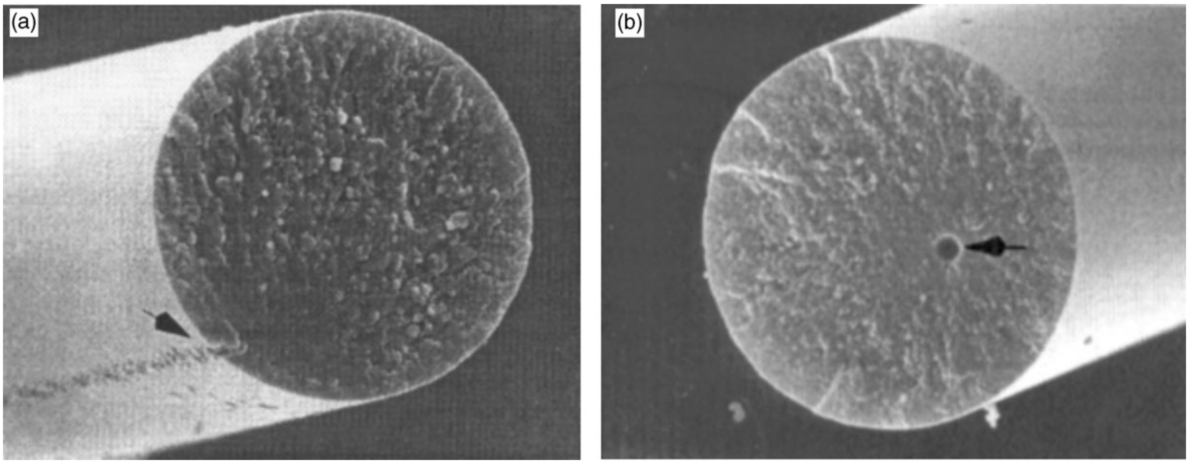


Figure 4
Fractographs of fibers with (a) linear surface defect and (b) spherical pore defect.

4. Microstructure–Strength Relationships in Fibers

Fibers are brittle, inelastic materials. Therefore, fiber strength is governed by the Griffith relationship:

$$\sigma = K_{\mathrm{Ic}}/\sqrt{\pi c} \tag{9}$$

where K_{Ic} is the fracture toughness of the fiber and c is the flaw size. Fracture toughness of fibers is no higher than bulk materials; in fact, the opposite is more likely. For instance, analysis of mirror-mist-hackle radii of SiC fiber fracture surfaces found that fiber fracture toughness was lower than in bulk SiC ceramics (Matsuo *et al.* 1992). In alumina, fracture toughness decreases with decreasing grain size. Thus, the very high strength that can be achieved by reinforcing fibers can only occur in fibers with very small grain and, hence, minimal flaw sizes. The Griffith equation predicts that a material with a fracture toughness of 3 (as expected for SiC or Al_2O_3) can achieve a fracture strength of 3 GPa only when the critical flaw size, c, is $<0.3\,\mu m$. Thus, not only must defect sizes be kept low by careful fiber processing, high fiber strength requires a very small grain size. The grain size of commercial fibers is usually below 0.1 μm.

Fiber defects can be identified by SEM observation of fracture surfaces captured during fractography testing. Fiber fracture is carried out in a damping medium such as grease to avoid secondary fracture and loss of the origin of fracture. Flaws in fibers are generally classified as two types: surface or internal flaws. Surface flaws occur due to abrasion, impacted particles, or other damage during processing. Figure 4(a) shows a Nextel 610 fiber that fractured at a linear surface defect caused by interfilament welding. Common internal flaws include porosity or regions of large grains. Figure 4(b) shows a Nextel 610 fiber that fractured at an internal pore.

Bibliography

Chi Z, Chou T, Shen G 1984 Determination of single fibre strength distribution from fibre bundle testing. *J. Mater. Sci.* **19**, 3319–24

Coleman B D 1958 On the strength of classical fibres and fibre bundles. *J. Mech. Phys. Solids* **7**, 60–70

Hurst J B, Hong W S, Gambone M L, Porter J R 1998 ASTM single fiber room temperature test standard development. American Society of Mechanical Engineers, paper 98-GT-567

Lavaste V, Besson J, Bunsell A 1995 Statistical analysis of strength distribution of alumina based single fibres accounting for fibre diameter variations. *J. Mater. Sci.* **30**, 2042–8

Matsuo Y, Li J X, Kimura S 1992 The relationship between tensile strength and flaw-size of silicon carbide single fibers. *Adv. Compos. Mater.* **2**, 17–28

van der Zwaag S 1989 The concept of filament strength and the Weibull modulus. *J. Test Eval.* **17**, 292–8

Wagner H D 1989 Stochastic concepts in the study of size effects in the mechanical strength of highly oriented polymeric materials. *J. Polym. Sci. B: Polym. Phys.* **27**, 115–49

Wilson D M 1997 Statistical tensile strength of Nextel 610 and Nextel 720 fibres. *J. Mater. Sci.* **32**, 2535–42

Zhu Y, Blumenthal W R, Taylor S T, Lowe T C, Zhou B L 1997 Analysis of size dependence of ceramic fiber and whisker strength. *J. Am. Ceram. Soc.* **801**, 447–52

D. M. Wilson
3M Co., St. Paul, Minnesota, USA

Fiberglass

Glass fibers have valuable properties as a result of their geometric shape and the uniqueness of glass itself. The room-temperature properties of glass make it well suited for many applications but it is its unusual viscous characteristics in the molten state and its resistance to crystallization that allow glass to be so easily formed into many shapes and, in particular, drawn into fibers.

Most metals, metal oxides, and salts form an extremely low-viscosity, almost water-like, liquid when melted; they also tend to crystallize rapidly at a specific temperature upon cooling. Any attempt to draw such low-viscosity, molten materials into fibers is thwarted by the surface tension overcoming the low viscous forces and causing the fiber to break up into spherical droplets to minimize surface energy. A few substances tend to form high-viscosity liquids when melted and have attendant low crystallization rates; such melts become progressively more viscous as they cool and they solidify into a glass. Uniform fibers can be drawn from these higher viscosity liquids since there isn't time during fiber formation for the relatively high viscous forces to be overcome by the surface tension.

Stable inorganic glasses can be made from a number of oxides and salts, but the most useful glasses are oxides. Silica (SiO_2), B_2O_3, P_2O_5, and GeO_2 are examples of some of the most common glass-forming oxides. Although silica is a valuable glass in its own right for both high-temperature low-expansion ware and low-loss optical fibers, most commercial glasses are a compound of silica and other oxides. The amounts and types of glass-forming and nonglass-forming oxides added to silica generate a range of glasses with widely varying properties in both the molten and solid state.

This article covers the basics of the manufacturing processes for fiberglass, of which there are two major forms.

(i) *Continuous glass fibers* (also referred to as reinforcement or textile glass fibers). These are constant-diameter, continuous glass threads or filaments which are attenuated from the molten state, usually by winding on to a rotating drum. Such fibers are, typically, processed into multifilament strands, mats, or woven cloths for use in composite materials.

(ii) *Insulation wool fibers*. These are discontinuous filaments, of random length and a range of diameters, produced by the attenuation of molten glass streams in a high-velocity gas, where the interaction of the gas with the glass stream is a primary attenuating factor. The pack of random, wool-like fibers so produced usually undergoes further processing and is then widely used for various forms of heat and sound insulation.

1. Continuous Glass Fibers

1.1 Compositions, Properties, and Uses

The composition ranges of a number of glasses used to produce continuous fibers are given in Table 1 with some attendant typical properties. E-glasses, and similar compositions containing little or no boron or fluorine, are the commonly produced glasses for reinforcement applications. They have an excellent

Table 1
Composition ranges in wt.% and typical properties for glass fibers.

	E-Glass	S-Glass	C-Glass	A-Glass	AR-Glass	Wool
SiO_2	52–56	64–66	64–68	63–72	55–75	55–70
Al_2O_3	12–16	24–25	3–5	0–6	0–5	0–7
B_2O_3	5–10	4–6	0–6	0–8	3–12	
CaO	16–25	0–0.1	11–15	6–10	1–10	5–13
MgO	0–5	9.5–10	2–4	0–4		0–5
BaO			0–1			0–3
Li_2O					0–1.5	0–0.5
Na_2O+K_2O	0–2	0–0.2	7–10	14–16	11–21	13–20.5
TiO_2	0–1.5			0–0.6	0–12	0–0.5
ZrO_2					1–18	
Fe_2O_3	0–0.8		0–0.8	0–0.5	0–5	0.1–0.5
F_2	0–1			0–0.4	0–5	0–1.5
Density, $g\,cm^{-3}$	2.58	2.46	2.52	2.44	2.7	2.54
Refractive Index	1.558	1.521	1.533	1.538	1.562	1.53
Softening Point, °C	846	1056	750	705	773	675
Tensile Strength, MPa	3445	4890	3310	3310	3241	3310
Young's Modulus, GPa	72.3	86.9	68.9	68.9	73.1	65.5
Coeff. Thermal Exp., $\times 10^{-7}\,°C^{-1}$	54	16	63	73	65	92

Compiled from data in Hartman *et al.* 1994, Eastes 1993, and unpublished sources.

combination of physical, electrical, and chemical properties. Glass fibers produced from these glasses are commonly found in fiber-reinforced plastics (FRP) used in electrical applications, automobiles, trucks, boats, pipes, and storage tanks. The other compositions in the table are used for more specialized applications.

S-glass is probably one of the more technically important glasses. It has the highest strength, stiffness, and softening point of any commercial reinforcement glass fiber. This glass is used in very demanding products such as FRP lightweight armor.

A-glass is a soda–lime–silica glass that is similar to common bottle glass and is sometimes used where mechanical performance and water durability are less critical.

C-glass is designed to have a high acid resistance and is used in such products as battery plate separators.

AR-glass has a high alkali resistance and was developed for use in the reinforcement of Portland cement.

1.2 Processing

Modern large-scale production of continuous glass filaments involves several steps (see Fig. 1). These include the weighing and mixing of batch materials (e.g., finely ground silica sand, clay, limestone, and colemanite or boric acid), the reaction and melting of the mixed batch at about 1500 °C, usually in a gas-fired or an electrically-boosted gas-fired furnace, the refining/homogenizing of the molten glass, the conditioning/cooling of the glass in the forehearth, and the subsequent drawing of fibers at about 1250 °C.

The actual formation of the fibers occurs through a platinum alloy "bushing," which is an electrically heated container situated below the forehearth of the furnace. The base of the platinum alloy bushing contains hundreds of short hollow tips arranged in rows, these tips having bores of about 1–2 mm in diameter. Molten glass at a viscosity of approximately 1000 P (100 Pas) exudes from each tip and is attenuated into fibers by winding on a rapidly rotating drum at fiber drawing speeds typically between 20 and 70 ms^{-1}. The diameter of the fibers is determined by the glass viscosity, tip bore dimensions, depth of the glass above the tips, and drawing speed. Usual commercial diameters range from 5 μm to 20 μm. Chilled metal fins are placed between the rows of tips creating a uniform environment for the tips and greater process consistency. Throughputs for a single bushing may be as high as 100 K gh^{-1} with several tens of bushings being supplied from a single furnace.

Water-based coatings (or "sizes") are usually applied to the fibers with a roller–applicator situated between the bushing and the winder. The applicator is immediately preceded by a water spray or mist which cools the fibers to prevent the chemical size from being applied to a hot glass surface. The coating on the fibers is subsequently dried and cured when the fiber package is removed from the winder and passed through an oven. The coating generally serves two functions: it protects the fibers from mechanical damage during the various process steps, and it also controls the interface properties between the fibers and the polymer matrix when they are formed into a composite material. Because of this, coatings are usually complex mixtures which invariably contain a coupling agent, in the form of a silane, to maintain the mechanical performance of

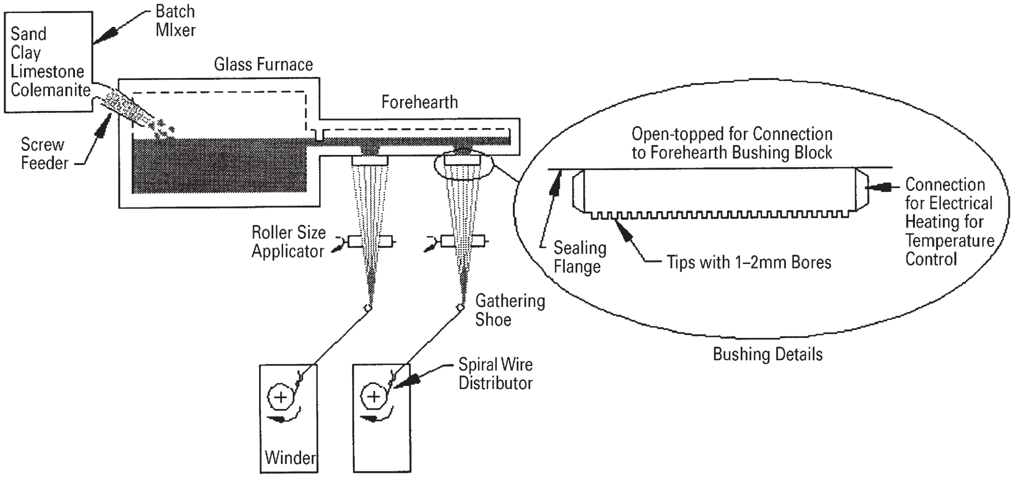


Figure 1
Continuous glass fiber manufacturing process.

subsequently made fiber/polymer composites in wet conditions.

Some specialty fibers are manufactured on a smaller scale using a resistively-heated, platinum-alloy melter situated directly above each fiberizing bushing to melt the batch materials. This process is commonly used for such glasses as S-glass. Other processes in use include those where the glass is made in a large furnace and then formed into marbles, which are subsequently remelted in small platinum alloy melters situated directly above each fiberizing bushing.

2. *Insulation Wool Fibers*

2.1 *Composition, Properties, and Uses*

Most commercial insulation glasses fall within the range depicted in Table 1. They invariably contain significant amounts of sodium and boron oxides to lower both the liquidus and working temperature, the latter to a level compatible with the operating temperature of the base metal spinners. An important requirement is that the fibers have good water durability to ensure that they do not deteriorate in shipping or service, requiring a balance between the amount and type of oxides needed for lower working temperature and those giving improved weatherability.

Insulation wool glasses find global use in commercial and residential buildings in both thermal and acoustical insulation as well as in the area of filtration.

Other important insulating fibers like rockwool, slagwool, and refractory ceramic fibers are, technically, glass fibers, too; the description of their processing is omitted here but can be found elsewhere (Eastes 1993).

2.2 *Processing*

In typical fiberglass insulation production the reaction and melting of the batch raw materials are carried out at 1400 °C to 1500 °C either in a gas-fired furnace or a submerged-electrode "cold-top" furnace. The batch ingredients, ground minerals such as silica sand, limestone, soda-ash, borax, and, increasingly, crushed post-consumer bottle and plate glass, are weighed, premixed, and fed into the furnace by a belt conveyor or a screw-feeder (see Fig. 2). Firing arrangements in the furnace coupled with stirring using bubblers and appropriate dwell times create a homogenous melt which is cooled to close to the working temperature in a forehearth channel before being drawn off as a glass stream. The primary division of this glass stream occurs while it is at a viscosity of about 1000 poise; the glass flows through several thousand holes in the vertical wall of a rotating bowl called a spinner (see inset in Fig. 2). The thousands of coarse primary fibers so produced are immediately attenuated further to the desired fineness by an annular external burner and/or a high-velocity turbulent gas stream which surrounds the spinner; typical final nominal fiber diameters range from 4 μm to 15 μm. A furnace may supply glass to as many as 10 spinners, with the throughput of each spinner being $450\,\mathrm{Kg\,h^{-1}}$ or more. The spinners operate at temperatures as high as 1050 °C and have to withstand hoop stresses of about 27 MPa with minimal distortion or creep in a severely corrosive environment.

The fibers produced by the above process are sometimes used as-made but are more commonly bonded together *en masse* by being sprayed with a water-based resin binder, typically phenol–formaldehyde (PF) based, which is subsequently dried and cured in an oven to give coherence to the fibrous material to allow for handleability. Continuous slitting and high-speed guillotining cuts the moving pack into appropriate dimensions for in-line compression packaging as rolls or batts. The importance of compression packaging, for economy in storing and shipping, may be appreciated when it is realized that

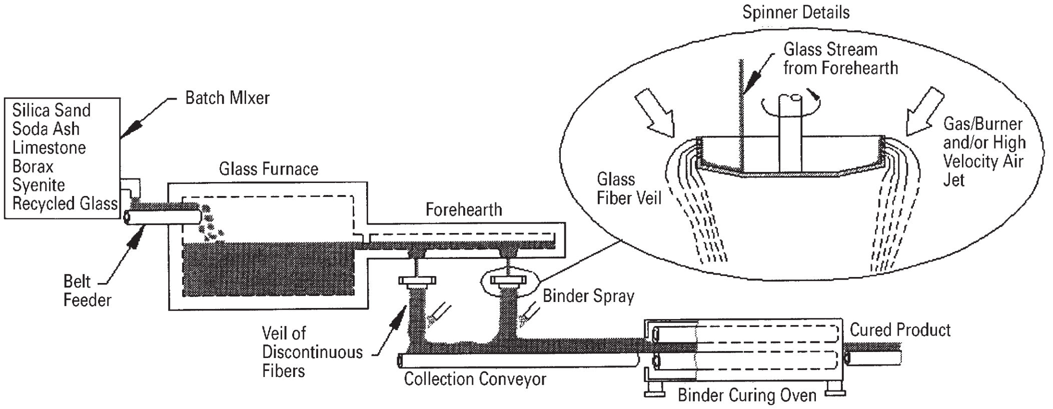


Figure 2
Insulation fiberglass manufacturing process.

approximately 98% of the volume of a typical uncompressed insulation batt is air space.

A specialty fibrous wool product of some technical significance, which is used for aircraft insulation and filtration, is produced by a flame-blowing technique. For this, marbles that have been produced by a furnace melting system similar to that described earlier are remelted in a platinum bushing and drawn into coarse filaments about 1 mm in diameter which are fed into a horizontally directed, high-energy gas burner. The extremely fine fibers produced in the turbulent gas jet are collected on a sloping conveyor.

3. Recent Developments

3.1 Processing

The most significant broad change in technology in recent years is the ongoing conversion to a new method of firing furnaces. This has been largely driven by environmental concern over nitrogen oxide (NO_x) emissions. The new firing technique, which is also being used in other segments of the glass industry, uses pure oxygen instead of air in fossil fuel-fired systems. Dramatic reductions in NO_x emissions are achievable but, due to higher flame temperatures, refractory brick changes are necessary. In addition, electrostatic precipitators have to be added to the furnace exhaust to capture the increased particulate volatilization from the molten glass surface.

3.2 Products

A product innovation with unique process challenges appeared with the commercial introduction of Miraflex® fibers in the mid-1990s. These fibers, which have remarkable curl and tortuosity, bear some analogy with the bimetallic strip. Two glasses, with markedly different coefficients of thermal expansion, are fused into a single fiber at the point of fiberization. The process requires two adjacent glass melting furnaces and forehearths, each with a different glass composition, to enable two separate glass streams to be dropped into a single insulation wool spinner. A system of baffles keeps the glasses separate until they are brought together at the spinner holes where they fuse, are attenuated, and then solidify without significant mixing of the compositions. A backscatter electron micrograph of the cross-section of the fibers (see Fig. 3) shows the delineated boundary between the two compositions. An interesting feature of this unbonded product is the process simplification brought about by the lack of need for a traditional resin binder application system and curing oven in manufacture.

3.3 Stewardship

As a consequence of much research in the last two decades or so of the twentieth century, the desirability of having low biopersistence in potentially respirable fibers was recognized. Regulations were developed within the European Union to ensure that insulations which contained respirable fibers were tested in animals and their compositions certified to have low biopersistence. The global nature of the industry has meant that most companies have compositional guidelines to ensure that this facet of the technology is appropriately controlled. Process implications are minimal since control is simply effected by adjusting the input batch materials to control glass chemistry.

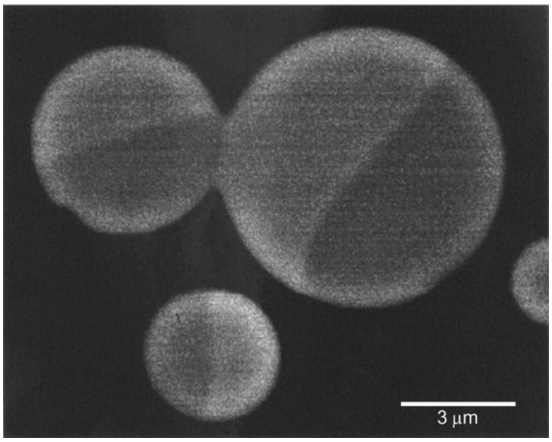


Figure 3
Backscatter scanning electron micrograph of Miraflex® fibers.

Bibliography

Aubourg P F, Wolf W W 1984 *Glass Fibers, Commercial Glasses*. Advances in Ceramics **18**. American Ceramic Society, Westerville, OH

Baukal C E (ed.) 1998 *Oxygen Enhanced Combustion*. CRC Press, Boca Raton, FL

Eastes W L (ed.) 1993 *Man-made Vitreous Fibers: Nomenclature, Chemistry and Physical Properties*. North American Insulation Manufacturers Association, Alexandria, VA

Gupta P K 1988 Glass fibers for composite materials. In: Bunsell A R (ed.) *Fiber Reinforcements for Composite Materials*. Elsevier, Amsterdam, pp. 19–71

Hartman D R, Greenwood M E, Miller D M 1994 *High Strength Glass Fibers*. Owens–Corning Fiberglass Corporation Pub. No. 1-PL-19025, Toledo, OH

Hook G E R, Lucier G W (eds.) 1994 Biopersistence of respirable synthetic fibers and minerals. In: *Environmental Health Perspectives Supplements*, Vol. 102, Supplement 5, US Dept. of Health, DHHS publication No. (NIH) 94-218

Houpt R A, Potter R M 1995 Dual-glass fibers and insulation products therefrom. US Pat. 5 431 992

Loewenstein K L 1993 *The Manufacturing Technology of Continuous Glass Fibres*. Elsevier, Amsterdam

Mikesell S L, Wolf W W 1986 Glass Fibers. In: Bever M B (ed.) *Encyclopedia of Materials Science and Engineering*. Pergamon, Oxford, pp. 1974–80

Mohr J G, Rowe W P 1978 *Fiber Glass*. Van Nostrand Reinhold, New York
Shand E B 1984 *Glass Engineering Handbook*. McGraw-Hill, New York
Tooley F V (ed.) 1984 *The Handbook of Glass Manufacture*. Ashlee, New York
Rawson H 1980 *Properties and Applications of Glass*. Elsevier, Amsterdam

N. M. Cameron and C. F. Rapp
Owens Corning Science and Technology Center,
Granville, Ohio, USA

Fibers: Superabsorbent

Absorbent fibers such as cotton, rayon, and wood pulp, have been used for a long time in articles ranging from diapers to surgical sponges. The transport of water in these cellulosic fibers is accomplished mainly by capillarity without significant swelling of the fibers. Since the 1970s, superabsorbent fibers have been developed that absorb many times the amount of water taken up by the earlier cellulosic fibers. These newer fibers with their unique structure swell on absorbing large quantities of water and become hydrogels capable of retaining most of the water even when they are exposed to pressure. The early superabsorbent fibers were produced by esterification or etherification reactions of the hydroxyl groups of cellulosic fibers along with cross-linking. One of the most popular superabsorbent cellulosic fibers is sodium carboxymethylcellulose that has been cross-linked.

An example of one of the newer synthetic superabsorbent fibers is Lanseal F produced by the Japan Exlan Company (now Toyobo). Lanseal F has an acrylic core with a superabsorbent outer layer that is claimed to swell about 12 times its initial size when it contacts water. Another acrylic superabsorbent fiber called Fibersorb developed by Arco Chemical Company has been manufactured by DuPont. A third synthetic superabsorbent fiber is reported to be a joint development of Allied Colloids and Courtaulds Fibers in the UK.

1. Mechanism of Water Absorption by Superabsorbers

Superabsorbent polymers are materials which, in the presence of water, exist in a gel state and as the name implies absorb large quantities of liquid. Although the definition is not strict, polymers are considered to be superabsorbent if they imbibe 20 times or more of their weight of aqueous fluid. Many are water soluble before being cross-linked, and the greatest uptake of water is usually at the lowest cross-linking density that provides water insolubility. This point is illustrated by Fig. 1, a plot of water absorbency vs. amount of cross-linking for poly(sodium acrylate).

The transport of water through a superabsorbent is complex and believed to occur by both capillary and diffusion mechanisms. Initial transport, when the water first contacts the superabsorbent, is primarily by capillarity. As a result of the initial sorption of water, the surface layer develops into a slowly permeable superabsorbent gel layer, which impedes the diffusion of water into the interior of unswollen polymer mass; this phenomenon is known as gel blocking. Cross-linking reduces gel blocking by preventing excessive surface swelling (Chen *et al.* 1985).

The water absorbency of polyelectrolyte superabsorbers is affected by the concentration of electrolytic salts in the water and the pH, as shown in Fig. 2 for sodium carboxymethylated cotton cross-linked in the

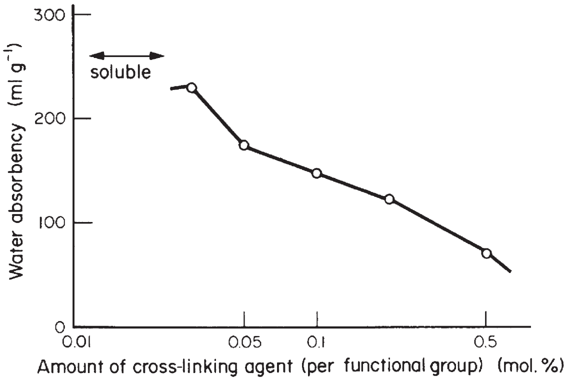


Figure 1
Water absorbency vs. cross-linking (after Chen *et al.* 1985; also Nukushin 1980) for poly (sodium acrylate).

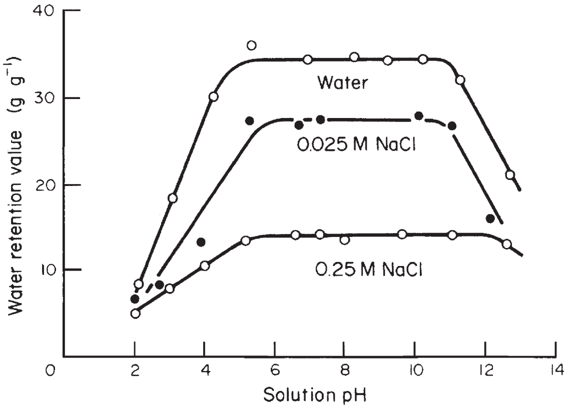


Figure 2
Water retention of cross-linked sodium carboxymethylated cotton (after Grignon and Scallan 1980).

wet state with epichlorohydrin. The swelling increases with pH to that characteristic of a weak acid, reaches a maximum near neutrality, then decreases at a pH of about 10. The whole curve is lowered as the concentration of added salt is increased. The effects of both pH and salt concentration are, of course, important in such practical uses of superabsorbents as diapers and incontinence products.

According to (Chen *et al.* 1985), superabsorbent polyelectrolyte gels generally have the greatest ability to absorb water when the pH of the water is between 8 and 10. The neutralization "acts to unravel and straighten out the polymer chains in aqueous fluids so that the final absorbent composition is more swellable in the presence of such fluids." Neutralization of ionic groups can be accomplished with a suitable base, including alkali metal hydroxides such as sodium hydroxide.

2. Cellulosic Fibers

Superabsorbent fibers may be prepared from cellulose by chemically modifying cotton, rayon, or wood fibers, or by extruding chemically modified cellulose. Cellulose ethers are the predominant cellulosic fiber superabsorbers. Methyl, ethyl, and carboxymethyl cellulose fibers are usually prepared from cellulosic fibers under alkaline conditions by treatment with alkyl chlorides. The cellulose ether can then be cross-linked to form a superabsorber. Alternatively, cross-linking may be accomplished prior to formation of the cellulose ether or even during formation. Cross-linking is normally carried out in the wet state because the fibers are then swollen and tend to maintain this swollen state with its high moisture absorbency.

Carboxymethyl cellulose (CMC) can be prepared by treating alkaline cellulosic fibers with monochloracetic acid. The degree of substitution (DS) achieved in this reaction is of critical importance. Although absorbency increases rapidly with an increase in DS, at a DS of about 0.35, the CMC fiber starts to become water soluble. Various methods have been employed to preserve the water insolubility of the fiber, including (i) cross-linking, (ii) heat treatment, and (iii) conversion of the ionic state (from metal salt to free acid). The increase in absorbency and water solubility due to an increase in DS has to be balanced by the decrease in solubility produced by any one or a combination of these methods so as to obtain optimum absorbency with sufficient insolubility to preserve the fibrous state of the superabsorbent fiber.

Other cellulose ethers that have provided superabsorbents are hydroxyethyl cellulose, carboxymethyl hydroxyethyl cellulose, and methylhydroxyethyl cellulose. Still another is phosphonomethylcellulose. In order to cross-link cellulose, a chemical must contain groups that will react with the hydroxy groups in cellulose and have a functionality of at least two. Formaldehyde, epichlorohydrin, and dichloro-2-propanol are commonly used cross-linkers.

Reid (1968) describes heat treatment of modified polysaccharides including cellulose ethers containing both carboxyl and hydroxyl groups capable of inter-reacting to form an ester cross-link at elevated temperatures. CMC is cross-linked by heating it at a low pH where there are free carboxyl groups to react with hydroxyl groups. The resulting absorbent CMC with a DS of about 0.7 and 65–70% of the carboxyl groups in the sodium salt form has good water absorbent properties. Hercules marketed a self-cross-linked CMC absorbent fibrous product based on Reid's work. The Aqualon subsidiary of Hercules is producing two such fibers called Aquasorb FC and FR, the latter being more highly cross-linked (Hercules 1986).

In addition to the superabsorbent fibers formed by cross-linking cellulose chains by esterification using a heat treatment, superabsorbent cellulosic fibers containing phosphorus have been prepared by the urea phosphate process, in which the fibers were treated with an aqueous solution of urea and orthophosphoric acid. Another cellulose ester, cellulose sulfate, has been used as a superabsorber as well.

Superabsorbent cellulosic fibers can be prepared by grafting onto cellulosic fibers hydrophilic groups or groups capable of being converted to hydrophilic groups after grafting. The former method, using acrylic and methacrylic acids, for example, has not produced fibers as highly absorbent as the latter method, which is illustrated by the grafting of acrylonitrile followed by saponification to sodium polyacrylate.

Takebe and Yamazaki (1978) disclosed the formation of highly absorbent fibers by spinning lower alkyl or lower hydroxyalkyl substituted cellulose ethers by wet or dry spinning methods. They spun fibers of methyl cellulose, ethyl cellulose, hydroxyethyl cellulose, hydroxypropyl cellulose, and hydroxypropylmethyl cellulose.

Superabsorbent cellulose ester fibers have also been made by wet spinning the maleic, succinic, and phthalic acid monoesters of cellulose (Diamantoglou and Meyer 1988). The alkali metal salts of the resulting fibers have superabsorbent properties.

3. Synthetic Superabsorbent Fibers

Burrow and Laity (1987) describe superabsorbent cross-linked polysaccharide fibers. These fibers can be produced by dissolving a carboxylate ester of a polysaccharide, such as a mixture of acetylated and benzoylated dextran, in a solvent, such as *N*-methyl pyrrolidone, with a latent cross-linking agent, such as epichlorohydrin, extruding the solution through a jet into an aqueous medium, and subjecting the fibers to conditions leading to hydrolysis of the ester groups and cross-linking of the resulting hydroxyl groups.

Gross (1976) spun superabsorbent fibers from an organic solvent solution of a copolymer of maleic anhydride and a vinyl aromatic monomer, such as styrene, with a cross-linking agent, such as an adduct of glycerine with ethylene oxide and propylene oxide, by dry spinning or wet spinning into a water bath. The resulting fibers were heated at an elevated temperature to complete the cross-linking and then were made water swellable by reaction with ammonia. Similarly, Bashaw and Atkins (1976) prepared superabsorbent fibers with copolymers of maleic anhydride and a vinyl monomer but used a polyfunctional cross-linking agent, such as a glycol or alkylene diamine, that will react with a carboxylic acid anhydride; the cross-linking was completed by heating at an elevated temperature followed by treatment with ammonia or an alkali metal hydroxide.

Holst *et al.* (1982) prepared superabsorbent fibers by hydrolyzing polyacrylonitrile or copolymer fibers at temperatures exceeding 100 °C in alkaline aqueous solutions, which could contain an organic solvent miscible with water.

Tanaka (1982, 1983) of the Toyobo Company has issued two patents that claim superabsorbent fibers containing an outer cross-linked hydrogel layer and an inner acrylonitrile core, and a process for producing them. The outer superabsorbent hydrogel layer is produced by exposing acrylonitrile fibers to an alkali metal hydroxide or ammonia. It is believed that Lanseal F, Toyobo's superabsorbent fiber, is covered by these patents. In the first, US Patent No. 4,366,206, fibers of polyacrylonitrile or acrylonitrile copolymer are partially hydrolyzed in an aqueous medium with (i) a high concentration of alkali metal hydroxide or ammonia, or (ii) a lower concentration of alkali metal hydroxide or ammonia in conjunction with an "electrolytic salt." The hydrolysis is carried out "… for a sufficient time to cross-link and render the outer layer of the fibers hydrophilic, thereby producing fibers having said hydrophilic cross-linked outer layer and a non-cross-linked, nonhydrophilic acrylonitrile inner layer." The patent claims fibers with less than 55% by volume of the outer hydrogel layer; it is preferred that the hydrogel layer is between 5% and 40% by volume. Although the patent claims fibers with water absorbency of up to 300 cm g^{-1}, the fibers with useful physical properties given in the examples have lower water absorbency values. The second patent, US Patent No. 4,374,175, is similar to the first but in addition claims superabsorbent fibers with "latent or visualized crimps" and includes a final step of wet heating the fibers at temperatures of 100–250 °C. This patent also claims "… an industrially advantageous process for producing …" the fibers.

Lanseal F is described as a "… double layer structure, with the outer layer of the superabsorbent substance … and an … inner layer of acrylic fiber …" (Toyobo Company). The rate of water absorbency and the absorbency in electrolyte solutions are given in Fig. 3. It is produced in two forms, a 5 denier fiber 51 mm long, and a 3 denier fiber 3 mm long. Lanseal F fibers, alone or in blends with other fibers, are recommended for a wide range of applications including feminine care products and disposable diapers.

Le-Khac (1988a, 1988b, 1989, 1990) of Arco Chemical Company has patented several superabsorbent fiber compositions. These patents cover superabsorbent fibers containing a copolymer of an α,β-unsaturated monomer, such as maleic anhydride, a suitable co-monomer, such as isobutylene, and a material capable of cross-linking the copolymer by means of a curing step at an elevated temperature after the fiber is spun. The preferred copolymer has 20–80% pendant carboxylic acid groups and 80–20% pendant carboxylate salt units from partial neutralization with a metal hydroxide. Fibers are formed from aqueous solutions by dry spinning. The cross-linking agents include compounds with at least two hydroxyl groups, which supposedly cross-link by reacting with the carboxylic acid groups of the copolymer, alkylene glycols with 2–10 carbons, for example, such as glycerol and its ether; other cross-linking agents include disaccharides such as sucrose and oligosaccharides.

Le-Khac (1988c, 1989, 1990) also patented superabsorbent fibers made from copolymers similar to those just described, but in place of the agents used as

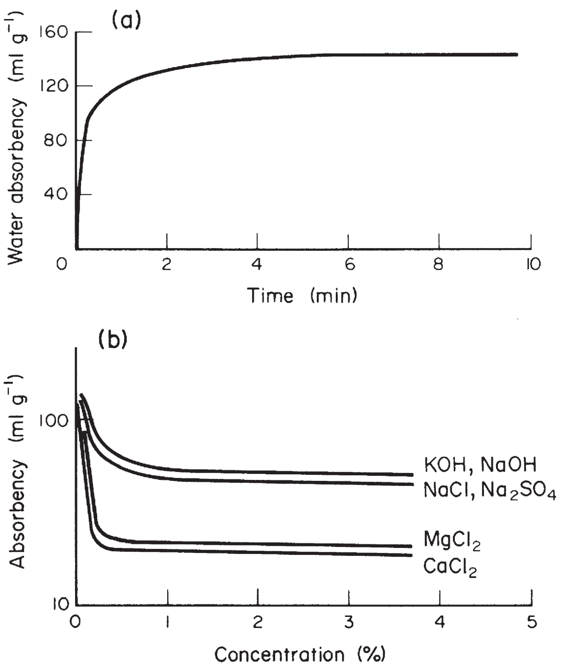


Figure 3
Absorbency behavior of Lanseal **F** fibers after Toyobo Company; (a) rate of water absorbency and (b) absorbency in electrolyte solutions.

cross-linkers he claimed polyethers derived from C_2 to C_{10} alkylene oxides "... wherein ... at least a portion of the oxide groups on the polyether are inter-reacted with the carboxyl hydrogen atoms on the pendant carboxylic acid units such that the copolymer and the polyether are associated through hydrogen bonds."

The composition of the superabsorbent fiber Fibersorb, developed by Arco Chemical Company, is unknown, but it may be covered by one of the Arco patents mentioned above. It is manufactured in three grades of fibers (Arco Chemical Company). The rate of water absorption of one of the fibers, SA-7000M, and its absorption under various pressures in water and in brine are shown in Fig. 4. Early in 1991, Arco Chemical Company announced that it was discontinuing Fibersorb development and offering to sell the Fibersorb operations.

There is little literature on another synthetic superabsorbent fiber being developed jointly by Allied Colloids and Courtaulds Fibers in the UK. The fiber is reported to be based on acrylate technology but uses a polymer system different from that of any other superabsorbent fiber. The fiber is said to have a saline absorptive capacity of up to 40 times its own weight. Allied Colloids' European Patent 0-268-498 describes superabsorbent polymers that can be made into fibers. The patent covers a substantially linear polymer formed by polymerization of a water-soluble ethylenically unsaturated monomer blend. This polymer contains both carboxylic acid and hydroxyl groups which cross-link to form an ester. The new fiber being developed by Allied Colloids and Courtaulds could possibly be based on this patent.

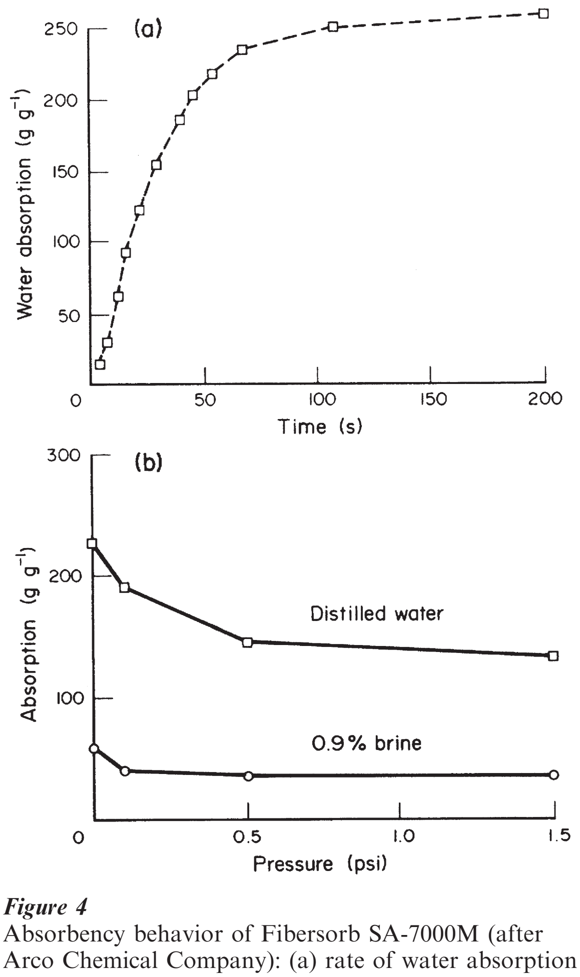


Figure 4
Absorbency behavior of Fibersorb SA-7000M (after Arco Chemical Company): (a) rate of water absorption and (b) absorption under various pressures.

4. Uses

Superabsorbent fibers have a wide range of uses. The more important applications are in products for sanitary purposes—in feminine care products, baby diapers, surgical dressings, hospital disposables, and incontinence fabrics for adults. They may be used alone but often are blended with other fibers not only to get the most favorable properties but also to economize, since superabsorbent fibers are usually more expensive. For example, blends with cellulose fluff give improved rates of absorption as well as more economical products. Superabsorbent fibers may be used in fiber form or in fabric structures containing superabsorbent polymer particles. One of the advantages of using fibers over polymer particles is that they usually migrate less in structures like diapers.

Superabsorbent fibers are used for the prevention of condensation; for example, Lanseal F fibers are used in a nonwoven fabric called Dew Stopper to absorb moisture in containers that carry coffee beans. They may be used to absorb water in oil and in organic solvents, although they will not absorb dissolved water. They are used in fabric form for gardening and agriculture. They may be used in fire prevention or in fire extinguishing operations. They also are used in filters for absorbing water. Attempts are being made to use superabsorbents for absorbing water in household goods, food products, pharmaceutical products, and medical devices.

Other uses for superabsorbent fibers will doubtlessly be developed in the future. Future improvements in absorbency and reduced costs of superabsorbent fibers might well open up for these fibers the large market now enjoyed by particulate superabsorbers.

See also: Textile Fibers: Mechanical Properties*; Textile Fibers: A Comparative Overview*

Bibliography

Arco Chemical Company Fibersorb® SA-7000 Superabsorbent Fibers. Technical Bulletin No. ACC OP51 7881 M. Arco Chemical Company, Newtown Square, PA

Bashaw R N, Atkins B L 1976 Absorbent fibers and process for their preparation. *US Pat.* 3 983 095

Burrow T R, Laity P R 1987 Absorbent fibers and their production. *Eur. Pat. Appl.* 0-232-121

Chen C, Vassallo J C, Chatterjee P K 1985 Synthetic and natural polymers. In: Chatterjee P K (ed.) 1985 *Absorbency*. Elsevier, Amsterdam, pp. 197–216

Diamantoglou M, Meyer G 1988 Process for the production of water-insoluble fibers of cellulose monoesters of maleic acid, succinic acid and phthalic acid, having an extremely high absorbability for water and physiological liquids. *US Pat.* 4 734 239

Grignon J, Scallan A M 1980 *J. Appl. Polym. Sci.* 25, 2829

Gross J R 1976 Absorbent fibers and process for their preparation. *US Pat.* 3 954 721

Hercules 1986 Aquasorb F Absorbent Fiber. Technical Bulletin No. 461-4. Hercules, Wilmington, DE

Holst A, Perplies E, Buchberger G 1982 Acrylic polymers having a high swellability in water. *US Pat.* 4 337 328

Le-Khac B 1988a Extended shelf life water-absorbing composition which facilitates fiber formation. *US Pat.* 4 731 067

Le-Khac B 1988b Sugar-containing water-absorbing composition which facilitates fiber formation. *US Pat.* 4 788 237

Le-Khac B 1988c Water absorbing polymer compositions and articles prepared therefrom. *US Pat.* 4 743 244

Le-Khac B 1989 Ultrahigh water-absorbing fiber forming composition. *US Pat.* 4 813 945

Le-Khac B 1990 Ultrahigh water-absorbing fiber-forming composition. *US Pat.* 4 892 533

Nukushin K 1980 Super absorbent. *Yuki Gosei Kagaku Kyokaishi* **38**, 546 (*Chem. Abstr.* 94, 66783e)

Reid A R 1968 Water-soluble polymers and process of preparing. *US Pat.* 3 379 720

Takebe T, Yamazaki T 1978 Water dispersible absorber for blood and the like. *US Pat.* 4 084 591

Tanaka K 1982 Novel water-swellable fibers having a high degree of water swellability and excellent physical properties and process for producing the same. *US Pat.* 4 366 206

Tanaka K 1983 Novel water-swellable fibers and process for producing the same. *US Pat.* 4 374 175

Toyobo Company Lanseal F, Super Absorbent Fiber. Technical Bulletin. Toyobo Company, Osaka, Japan

T. F. Cooke
TRI, Princeton, New Jersey, USA

Fibrous Reinforcements for Composites: Overview

The fibrous form of matter is very common in nature. This is because matter in this form is highly flexible and thus can be used to make complex and multifunctional parts (Chawla 1998a). Modern usage of synthetic fibers as reinforcement for polymers, metals, and ceramics to make composite materials can be traced to the use of glass fibers for reinforcement of polyester and epoxy thermoset matrix materials in the first half of the twentieth century. High-stiffness and high-strength fibers (organic and inorganic) came about in the second half of the twentieth century, more particularly in the last quarter of the twentieth century. This can be attributed to advances in synthesis and processing as well as to an understanding of the importance of a processing–structure–property triad. This triad implies that the processing of a material into a fibrous form determines its microstructure, which determines the ultimate properties of the fiber.

1. A Brief Historical Perspective

In the late 1930s and early 1940s it was discovered that silica-based glass could be drawn into a very high-strength fiber. The stiffness of the glass fiber does not differ from that of the bulk glass. This is because glass is an amorphous material, and therefore there is no preferential orientation after the fiber-drawing operation. The stiffness of glass is, however, quite high compared to that of most polymers and therefore glass fiber is used as a reinforcement fiber for polymeric materials.

The last quarter of the twentieth century saw a large body of work that led to the commercial production of high-stiffness and high-strength fibers, organic as well as inorganic. In the 1950s it was realized that the carbon–carbon bond in the backbone chain of polymers is very strong and that if one could only orient the molecular chains, one should get a high-modulus fiber from the organic fibers. This was attempted at first by applying ever-increasing stretch or draw ratios to the spun organic fibers. This resulted in some improvement in the stiffness of the fiber, but no better than was available with glass fibers, i.e., a tensile or Young's modulus of about 70 GPa in the fiber direction. It was at this stage that a major structure–property relationship in polymers was realized, namely that one needs orientation *and* extension of the molecular chains in order to realize the full potential of the carbon–carbon bond. Aramid and ultrahigh molecular weight polyethylene fibers, with a Young's modulus over 100 GPa, epitomize this oriented and extended chain structure.

The last quarter of the twentieth century also saw an increasing amount of work in the area of ceramic fibers having low density, high stiffness, and high strength, but more importantly possessing these characteristics at high temperatures (as high as 1000 °C and higher). Examples of such high-temperature fibers include boron, carbon, alumina, silicon carbide, etc. Some very novel and innovative techniques based on the sol–gel process and the use of polymeric precursors to obtain inorganic or ceramic fibers such as silicon carbide, alumina, etc., came into being. The driving force for this development was the use of these fibers as reinforcements for metals and ceramics, i.e., at medium to very high temperatures.

2. Stiff and Strong Fibers

The use of fibers as high-performance materials in structural engineering applications is based on three important characteristics:

(i) *High strength*. This is achieved through several mechanisms. Because materials in fibrous form feature a small diameter with respect to the grain size or other microstructural units of the material, in some systems this allows a higher fraction of the theoretical strength to be attained than is possible with the same material in bulk form. This is also known the size effect, i.e., the smaller the size, the lower the probability of having an imperfection of a critical size that would lead to the failure of the material. Then, for a material in its fibrous form, the strength decreases as its diameter increases. In highly oriented carbon or polymeric fibers, as mentioned above, orientation of atomic bonds in the material along the fiber axis results in a fibrous material having far higher strength and, sometimes, even a higher elastic modulus than in the bulk form. In some fibers, particularly when produced via the sol–gel route or by conversion of preceramic polymers, high strength is achieved because microstructural units of the material, such as grain size and with which strength-limiting defect dimensions are scaled, can be produced with a particularly small dimension. In fibrous form, then, these materials are intrinsically stronger.

(ii) *A very high degree of flexibility*. This is really a characteristic of a material having a high modulus and a small diameter. Even materials that are brittle in their bulk form such as glass, alumina, silicon carbide, etc., when given the form of a fine-diameter fiber, can become quite flexible. This flexibility allows a variety of techniques to be employed for producing yarns from continuous fibers, which in turn can be braided, knitted, or woven into fabrics, ropes, cords, and other complex shapes and forms for use as reinforcement in a matrix material.

(iii) *A high aspect ratio (length/diameter, R/d)*. This allows a very large fraction of applied load to be transferred via the matrix to the stiff and strong fiber in a fiber-reinforced composite (e.g., Chawla 1998b).

Thus, in summary, the most distinctive feature of a fibrous material is that its properties are highly biased along its length. Fibers in general, and continuous fibers in particular, are hence very attractive reinforcements for composite materials. These characteristics have led to the development of fiber-reinforced composites with a variety of matrix materials such as polymers, metals, glasses, and ceramics (e.g., Chawla 1993, 1998b, Clyne and Withers 1993, Hull and Clyne 1996, Piggott 1980, Taya and Arsenault 1989, Suresh *et al.* 1993).

3. Important Definitions

Several important terms, definitions, parameters, and features serve to characterize the wide variety of existing fibers and products made from fibers (other than fiber-reinforced composites). These definitions and features are generally independent of fiber type, i.e., polymeric, metallic, glass, or ceramic fibers. They depend on the geometry rather than any material characteristics.

A fiber can be defined as an elongated material having a more or less equiaxed and uniform transverse cross-section of diameter or thickness less than 250 μm, and an aspect ratio, defined as the ratio of fiber length to fiber cross-sectional diameter or thickness, greater than about 100. This definition is a purely geometrical one in that it applies to any material. Often, however, the aspect ratio falls below this value but the material is still, by virtue of its elongated cylindrical shape and fine diameter, clearly fibrous in form; in such cases, fibers are designated as short fibers, chopped fibers, whiskers, or staple fibers depending on their nature and geometry (see below).

A fiber is the fundamental unit in making textile yarns and fabrics. Fibers can be naturally occurring or synthetic. There are many natural fibers, organic and inorganic. Examples include organic fibers such as silk, wool, cotton, jute, sisal; and inorganic fibers such as asbestos and basalt. There are a large variety of synthetic fibers available commercially. Polymer fibers such as polypropylene, polyethylene, polyamides, poly(ethylene terephthalate) (PET), polyacrylonitrile (PAN), polytetrafluoroethylene (PTFE), aramid, etc., are well-established fibers. Metallic wires or filaments have been available for a long time. Examples include steel, aluminum, copper, tungsten, molybdenum, gold, silver, etc. Among ceramic and glass fibers, glass fiber for polymer reinforcement has been available since the 1940s; optical glass fiber for telecommunication purposes made its debut in the 1950s, while ceramic fibers such as carbon, silicon carbide, alumina, etc., became available from the 1960s.

One can transform practically any material, polymer, metal, or ceramic, into a fibrous form. Historically and traditionally, fibers were used in the textile industry for uses such as clothing, upholstery and draperies, sacks, ropes, cords, sails, containers, etc. Gradually, their use entered the realm of more engineered items such as conveyor belts, drive belts, geotextiles, etc. With the advent of high-modulus fibers, the use of fibers has extended to highly engineered materials such as composites. Understandably, therefore, the terms commonly used with fibers have their origin far more often in textile technology terminology than in terminology specifically developed for or by the materials community. Relevant definitions include:

Aspect ratio. The ratio of length to diameter of a fiber (see above).

Braiding. Two or more yarns are intertwined to form an elongated structure. The long direction is called the bias direction or machine direction.

Chopped strand. Fibers in a strand are chopped to various lengths, 3–50 mm, for mixing with resins.

Continuous fibers. Continuous strands of fibers, generally available as wound fiber spools.

Denier. A unit of linear density. It is the mass in grams of a 9 km long yarn. This unit is commonly used in the US textile industry.

Fabric. A kind of planar fibrous assembly. It allows the high degree of anisotropy characteristic of yarn to be minimized, although not completely eliminated.

Filament. Continuous fiber, i.e., fiber with an aspect ratio approaching infinity.

Fill. See weft.

Monofilament. A large-diameter continuous fiber, generally with a diameter greater than 100 μm.

Roving. A bundle of yarns or tows of continuous filaments (twisted or untwisted).

Spinneret. A vessel with numerous shaped holes at the bottom through which a material in a molten state is forced out in the form of fine filaments or threads.

Staple fiber. Fibers having short, discrete lengths (10–400 mm long) that can be spun into a yarn are called staple fibers. This spinning quality can be improved if the fiber is imparted a waviness or crimp. Staple fibers are excellent for providing bulkiness for filling, filtration, etc. Frequently, staple natural fibers, e.g., cotton or wool, are blended with staple synthetic fibers, e.g., nylon or polyester, to obtain the best of both types.

Tenacity. A measure of fiber strength that is commonly used in the textile industry. Commonly, the units are gram-force per denier, gram-force per tex, or newtons per tex. It is a specific strength unit, i.e., there is a factor of density involved.

Tex. A unit of linear density. It is the weight in grams of 1km of yarn. Tex is commonly used in Europe.

Tow. Bundle of twisted or untwisted continuous fibers. A tow may contain tens or hundreds of thousands of individual filaments.

Twist. The angle of twist that individual filaments may have about the yarn axis. Most yarns have filaments twisted because it is easier to handle a twisted yarn than an untwisted one.

Warp. Lengthwise yarn in a woven fabric.

Weft. Transverse yarn in a woven fabric. Also called *fill*.

Whisker. Tiny, whisker-like fiber (a few millimeters long, a few micrometers in diameter) that is a single crystal and almost free of dislocations. Note that this term involves a *material* requirement. The small size and crystalline perfection make whiskers extremely strong, approaching the theoretical strength.

Wire. Metallic filament.

Woven fabric. Flat sheet able to be draped made by interlacing yarns or tows.

Woven roving. Heavy fabric able to be draped woven from continuous rovings.

Yarn. A generic term for a bundle of untwisted or twisted fibers (short or continuous).

4. Fiber Flexibility

The flexibility of a fiber, i.e., its ability to be bent to an arbitrary radius, is one of the most important attributes of a fibrous material and worth discussing in some detail. It is easy to visualize that many operations such as weaving, braiding, winding, etc., depend on the ability of a fiber to be bent without breaking. This flexibility is also of great importance in composites because it permits a variety of techniques to be employed for making composites with these fibers. If we treat the fiber as an elastic beam of circular cross-section, then we can easily see that the fiber flexibility corresponds to the ability of the elastic beam to be bent to an arbitrary radius of curvature. A high degree of flexibility is really a characteristic of a material having a low modulus and a small diameter (Dresher 1969). The flexibility of a given fiber is a function of its stiffness or elastic modulus, E, and the moment of inertia of its cross-section, I.

The elastic modulus of a material is quite independent of its form or size. It is generally a material constant for a given chemical composition and fully dense material. Thus, for a given composition and density, the flexibility of a material is determined by its shape, the size of its cross-section, and its radius of curvature, the maximum attainable value of which is a function of the strength of the material. We can use the inverse of the product of bending moment (M) and the radius of curvature (R) as a measure of flexibility. From elementary strength of materials, we have the following relationship for a beam in bending:

$$M/I = E/R$$

Rearranging, and remembering that the moment of inertia, I, for a beam of circular cross-section of diameter d is equal to $\pi\ d^4/64$, we can write:

$$MR = EI = E\pi d^4/64$$

The product term EI ($= MR$) is known as flexural rigidity. So, we can define the flexibility of a fiber as:

$$1/MR = 64/E\pi d^4 \approx 20/Ed^4$$

This equation indicates that ($1/MR$), a measure of flexibility, is a very sensitive function of diameter, d. If we take a 25 μm diameter nylon fiber as a classical example of a flexible fiber, we can compute the diameter of various other fibers that will be required to have a flexibility equal to this 25 μm diameter nylon fiber, see Fig. 1. It follows from this curve that, given a sufficiently small diameter, it is possible to produce, in principle, an equally flexible fiber from a polymer,

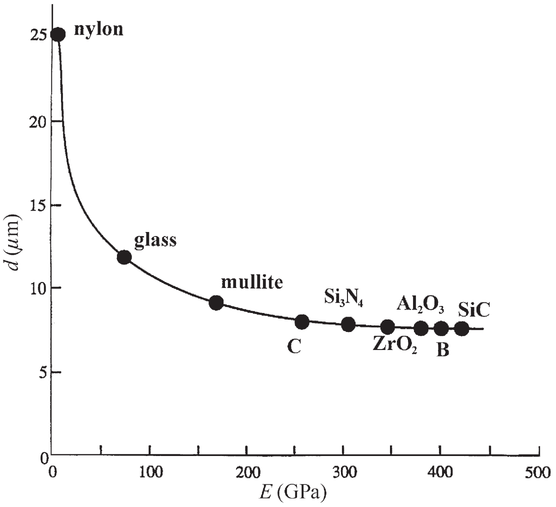


Figure 1
Diameter of various fibers with a flexibility equal to that of a 25 μm diameter nylon fiber. Given a sufficiently small diameter, it is possible to produce, in principle, an equally flexible fiber from a polymer, a metal, or a ceramic.

a metal, or a ceramic. In other words, one can make a very flexible fiber out of a ceramic such as silicon carbide or alumina provided one can produce it with a fine enough diameter.

5. *Fiber Processing*

It is possible to process a variety of materials into a fibrous form; in particular, tremendous progress has been made in processing high-stiffness and high-strength organic fibers (e.g., aramid, polyethylene) and inorganic fibers (e.g., silicon carbide, alumina

The process of producing a fibrous form from the liquid state is called spinning. There are quite a few variants of the basic fiber-spinning process. The most common is called melt spinning, which as the name implies means that the fiber is produced from a melt. This process is used for producing fibers from organic polymers and inorganic silica-based glasses. An appropriate thermoplastic polymer, generally in the form of dry pellets, is fed into an extruder where it is melted and then pumped through a spinneret to form filaments. A spinneret has multiple holes through which the molten material exits as fine filaments. Figure 2 shows a schematic of a generic fiber-spinning process. Suffice it to mention here three important types of polymeric fiber-spinning processes: dry spinning, wet spinning, and dry jet–wet spinning. In the dry spinning process the polymer melt or solution is extruded into an evaporating gaseous stream. In the wet-spinning process molten jets are extruded into a precipitating liquid medium. In dry jet–wet spinning the extruded solution passes through an air gap before entering a coagulation bath.

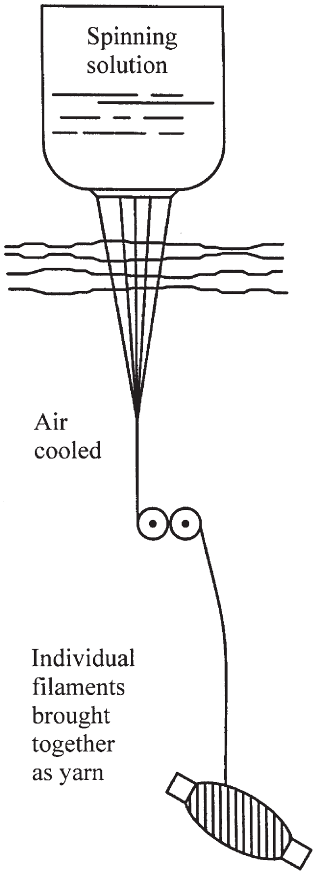


Figure 2
Schematic of a generic fiber-spinning process.

The following process variables are very important in any fiber-spinning process: (i) flow rate of the melt, (ii) take-up velocity of the wound filaments, (iii) extrusion temperature, (iv) dimensions and shape of the spinneret holes, and (v) cooling conditions.

The production of metallic fibers from the melt is generally difficult, mainly because of the high temperatures often involved and also because the high surface tension of molten metals implies a higher tendency towards destabilization of metallic melt streams. Metallic wires are, therefore, generally produced by taking advantage of their high ductility in the solid state, by drawing through a die. This very old process has the further advantage of leading to a microstructure featuring grains or second phases that are strongly oriented along the fiber axis, which generally confers on drawn metallic wires elevated tensile strengths along their axes.

Processing of carbon or of ceramics other than glass into fibrous form generally requires very high melt temperatures and the melts generally do not have the rheological characteristics suitable for the melt-drawing process. Thus, to make such fibers one must use other methods. One such method is chemical vapor deposition (CVD). Silicon carbide and boron monofilaments are typically made by CVD on a substrate heated to around 1300 °C (DeBolt *et al.* 1974). The substrate can be tungsten or carbon. The reactive gaseous mixture contains hydrogen and alkyl silanes for SiC or BCl_3 for boron and very high frequency (~60 MHz) to obtain an optimum temperature profile. To obtain a 100 μm SiC monofilament, it generally takes about 20 s in the reactor. The filament is wound on a spool at the bottom of the reactor.

The exhaust gases (95% of the original mixture + HCl) are passed around a condenser to recover the unused silanes. Efficient reclamation of the unused reactant gases is very important for a cost-effective production process. This process results in a fiber structure featuring elongated grains along the radial direction, at times producing a fiber surface that is somewhat nodular, the nodules on the surface of SiC being smaller than those seen on boron fibers. Such CVD processes result in composite monofilaments which have built-in residual stresses. The process is, of course, very expensive.

The SiC fiber obtained via CVD is very large in diameter, and thus not very flexible. An alternative route to obtain a fine, continuous, and flexible SiC fiber was developed in the mid-1970s by Yajima and co-workers (Yajima *et al.* 1976, Yajima 1980). They used a polycarbosilane precursor to make a nominally silicon carbide fiber, which is manufactured under the trade name of Nicalon. Actually, Nicalon fiber consists of SiC, SiO_2, and some free carbon. This method of using silicon-based polymers to produce a family of ceramic fibers having good mechanical properties, good thermal stability, and oxidation resistance has enormous potential. In any process involving ceramic fiber fabrication via a polymer route, the following are the critical steps (Wax 1985):

(i) polymer characterization (yield, molecular weight, purity, etc.);

(ii) melt-spin the polymer into a precursor fiber;

(iii) cure the precursor fiber to cross-link the molecular chains, making it infusible during the subsequent pyrolysis. This step may not be required under certain conditions, see below; and

(iv) controlled pyrolysis.

Carbon fiber is also often fabricated by controlled pyrolysis of an organic fiber precursor such as PAN. Depending on the precursor and processing, one can obtain a variety of carbon fibers with different strength and modulus. Important types are high strength (HS), high modulus (HM), intermediate modulus (IM), and super high modulus (SHM). There are certain essential steps common to all processes of carbon-fiber manufacture that use conversion of a polymer precursor.
These are:

(i) *Fiberization.* This involves converting the precursor material into a fibrous form, i.e., extrusion of a polymer melt or solution into a precursor fiber.

(ii) *Stabilization.* This treatment renders the precursor infusible during the subsequent high-temperature processing. This involves preoxidation or cross-linking of the polymer by thermosetting and is done at relatively low temperatures (200–450 °C), usually in air.

(iii) *Carbonization.* This step involves conversion of the stabilized precursor fiber into a carbon fiber. This is carried out in an inert atmosphere (generally pure N_2) at a temperature between 1000 °C and 2000 °C. At the end of this step the fiber has 85–95% carbon content.

(iv) *Graphitization.* This is an optional treatment. It is done in argon or N_2 at a temperature between 2400 °C and 3000 °C. This step increases the carbon content to more than 99% and imparts a very high degree of preferred orientation to the fiber.

Alternatively, carbon fibers can be produced from pitch-based precursors. Pitch is a generic term. Asphalt pitch is a product of oil refining while coal tar pitch can be obtained by mixing coal with appropriate organic fluids. Pitch is converted into a liquid crystal or mesophase pitch, which is a mixture of aromatic hydrocarbons with generally high molecular weights but with a distribution of molecular weights. The mesophase pitch is melt-spun into a fibrous precursor form. The pitch precursor fiber is then pyrolyzed under controlled conditions to obtain a highly oriented graphitic fiber. Pitch-based carbon fibers can have a very high elastic modulus and exceptionally high thermal conductivity.

Unlike glass fibers, many oxide ceramic fibers cannot be spun directly from the melt because the temperatures involved are far too high for the process to be economical (tooling and energetic requirements then become prohibitive). Therefore, to produce fibers of oxide ceramics such as alumina or mullite, one generally spins a slurry or other precursor of the ceramic at temperatures far below its melting temperature, after which the process comprises a sintering or firing step. A low firing or sintering temperature will result in a small grain size but is very likely to lead to an unacceptable level of residual porosity. At higher processing temperatures, one can eliminate porosity but excessive grain growth may result. Control of this step of the process is, thus, important in this fiber-processing route. Another approach for the production of refractory oxide fibers is the sol–gel route. Many companies have commercialized alumina or alumina-based fibers produced by sol–gel processing. An important series of such fibers, produced by 3M Co., goes by the name of the Nextel series of fibers.

Finally, we should mention an important form of short fiber, namely whiskers. These are monocrystalline, short fibers with extremely high strength. Strength levels approaching the theoretically expected values have been measured in many whiskers. Typically, whiskers have a diameter of a few micrometers and a length of a few millimeters. Thus, their aspect ratio (length/diameter) can be very high, varying from 50 to 10 000. This makes it attractive to use these high-strength whiskers as reinforcements. Whiskers do not have uniform dimensions or properties. This is perhaps their greatest disadvantage, i.e., the spread in properties is extremely large. It should be also recognized that preservation of a defect-free surface during processing of a composite is a very difficult task, i.e., handling and alignment of whiskers in a matrix to produce a composite are rather difficult. Whiskers are normally obtained by vapor phase growth.

Early in the 1970s a new process was developed, starting from rice hulls, to produce SiC particles and whiskers (e.g., Chawla 1998a). The SiC particles produced by this process are of a fine size. Rice hulls, a waste by-product of rice milling, contain cellulose, silica, and other organic and inorganic materials. Raw rice hulls are heated in the absence of oxygen at about 700 °C to drive out the volatile compounds. These coked rice hulls, containing about equal amounts of SiO_2 and free carbon, are heated in an inert or reducing atmosphere (flowing N_2 or NH_3 gas) at a temperature between 1500 °C and 1600 °C for about 1h to form silicon carbide:

$$3C + SiO_2 \rightarrow SiC + 2CO$$

When the above reaction is over, the residue is heated to 800 °C to remove any free carbon. Generally, both particles and whiskers are produced together with some excess free carbon. A wet process is used to separate the particles and the whiskers. Typically, the average aspect ratio of the as-produced whiskers is 75.

A vapor–liquid–solid (VLS) technique can be used to grow exceptionally strong and stiff silicon carbide whiskers. The catalyst forms a liquid solution interface with the growing crystalline phase while elements are fed from the vapor phase through the liquid–vapor interface. Whisker growth takes place by precipitation from the supersaturated liquid at the solid–liquid interface. The catalyst must take in solution the atomic species of the whisker to be grown. For SiC whiskers, transition metals and iron alloys meet this requirement.

6. Skin–Core Structure

Melt or solution spinning is one of the most common methods of giving a fibrous form to a material. Even in the case of ceramic fibers that are produced from polymeric precursor fibers, high-speed spinning is used to produce the precursor fiber. This high-speed spinning process can have some important effects on the resultant fiber structure and, consequently, on its properties. The process of extrusion through a spinneret leads to a certain chain orientation in the filament. Generally, the molecules in the surface region undergo more orientation than the ones in the interior because the solution in contact with the sides of the spinneret hole suffers more frictional resistance than the solution in the center of the spinneret hole. In other words, high-speed spun fibers show a skin–core structure, i.e., there are radial variations in the microstructure of a fiber (Fig. 3)

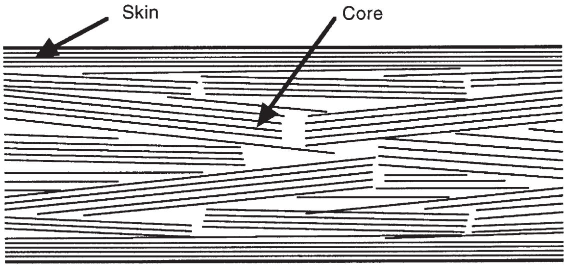


Figure 3
Schematic of a skin–core structure in a high-speed spun fiber showing a greater degree of crystallization in the skin than in the core.

Microstructurally, these variations correspond to a higher degree of molecular orientation and crystallinity in the near-surface region (skin region) than in the core region of the fiber (Perez 1985, Vassilatos *et al.* 1985). Another reason for this skin effect is the presence of a radial temperature variation (lower temperature in the near-surface region than in the interior) due possibly to cross flow of the quench air. Such a radial temperature variation results in a higher viscosity and higher stress in the skin region than in the core. Consequently, most fibers spun at high speed show a greater degree of crystallization in the skin than in the core. This skin effect can affect many other properties of the fiber, e.g., the adhesion with a polymeric matrix or the ability to be dyed.

7. Stretching and Orientation

As pointed out above, the spinning process of making a fiber results in some chain orientation but this occurs together with a skin effect. Generally, the as-spun polymeric fiber is subjected to some stretching, causing further chain orientation along the fiber axis and consequently better tensile properties, such as stiffness and strength along the fiber axis. The amount of stretch is generally described in terms of a draw ratio, which is the ratio of the initial diameter to the final diameter. The higher the draw ratio, the higher the degree of chain alignment, the higher the degree of crystallinity, and consequently the higher will be the elastic modulus of the fiber along its axis. This can affect a number of fiber characteristics, e.g., the ability of a fiber to absorb moisture: the higher the degree of crystallinity, the lower the moisture absorption. Du Pont produces a series of aramid fibers under the trade name of Kevlar. In particular, Kevlar 149 fiber has a higher degree of crystallinity than Kevlar 49. Consequently, Kevlar 149 has a higher elastic modulus and absorbs less moisture than Kevlar 49. In general, one can say that a higher degree of crystallinity in a polymeric fiber translates into a higher resistance to penetration by foreign molecules, i.e., a greater chemical stability.

There is, however, a limit to the amount of stretch that can be given to a polymer because the phenomenon of necking can intervene and cause rupture of the fiber. In other words, in a polymeric fiber there is a limit to the modulus enhancement that can be obtained by subjecting it to an ever-higher draw ratio.

8. *Cross-sectional Shape and Surface Roughness*

Although many fibers are more or less circular in cross-section, fibers can have variety of cross-sectional shapes. Natural fibers, in particular, mostly have noncircular cross-sections. For example, cotton fiber has a cross-section that appears like a collapsed circle. The cross-sectional shape of a spun synthetic fiber can be varied by changing the shape of the spinneret holes through which the polymer is extruded. In fact, one can even produce hollow fibers. Synthetic fibers can have dog bone, kidney bean, or trilobal cross-sectional shapes. Most melt-spun fibers, however, have initially a circular cross-section; however, solution spinning can still result in lobed or dog-bone-type cross-sections even though the spinneret orifice may be circular. The reason for this is that although the jet emerging from the orifice is circular, it collapses as the solvent is removed by drying (in dry spinning) or by solution (in wet spinning). Such noncircular fibers can provide a greater density in a fabric than circular fibers.

The surface texture of fibers can also vary considerably. Vegetable fibers such as cotton, jute, sisal, etc., have an irregular surface, while most animal fibers such as hair have scales. Compared to the natural fibers, synthetic fibers generally have relatively smooth surfaces; however, synthetic fibers can also show a large degree of variation in their surface texture. Striation markings due to drawing are commonly observed on the surface of metallic filaments. Similar markings, but on a finer scale, are observed on the surface of carbon fibers. As mentioned above, fibers produced by CVD show a somewhat nodular surface. Very fine-grained ceramic fibers can show surface roughness that scales with the grain size of the

Table 1
Properties of some important synthetic fibers.

Material (fiber)	Tensile modulus (GPa)	Tensile strength (GPa)	Compressive strength (GPa)	Density (gcm $^{-3}$)
Steel	200	2.8		7.8
Al alloy	71	0.6		2.7
Ti alloy	106	1.2		4.5
Alumina	350–380	1.7	6.9	3.9
Boron	415	3.5	5.9	2.5–2.6
SiC	200	2.8	3.1	2.8
S-glass	90	4.5	>1.1	2.46
Carbon P100 (pitch-based)	725	2.2	0.48	2.15
Carbon M60J (PAN-based)	585	3.8	1.67	1.94
Kevlar 49	125	3.5	0.39–0.48	1.45
Kevlar 149	185	3.4	0.32–0.46	1.47
PBZT	325	4.1	0.26–0.41	1.58
PBZO	360	5.7	0.2–0.4	1.58
Spectra 1000	172	3.0	0.17	1.0
Vectran	65	2.9		1.4
Technora	70	3.0		1.39
Nylon	6	1.0	0.1	1.14
Textile PET	12	1.2	0.09	1.39

fiber (Chawla *et al*. 1993). Such surface roughness characteristics can have an important bearing in regard to mechanical bonding of fibers with matrix materials in different composites (Chawla 1993).

9. Density

The density of fibers is a very important attribute. Mentioned above was the use of linear density or mass per unit length. That term is really suitable for giving an idea of the fiber size. Bulk density or mass per unit volume tells us how heavy a material is. Bulk density values of some important fibers are given in Table 1. It can be easily verified that if we divide the linear density of a fiber by its bulk density we get the cross-sectional area of the fiber. Most polymeric fibers have bulk densities varying between $0.9\,\mathrm{g\,cm^{-3}}$ and $1.5\,\mathrm{g\,cm^{-3}}$ (polypropylene, $0.9\,\mathrm{g\,cm^{-3}}$; polyethylene, $0.97\mathrm{g\,cm^{-3}}$; polyester, $1.4\,\mathrm{g\,cm^{-3}}$; cotton, $1.5\,\mathrm{g\,cm^{-3}}$). Ceramic fibers have densities varying between $2\,\mathrm{g\,cm^{-3}}$ and $4\,\mathrm{g\,cm^{-3}}$ while metallic fibers can have densities in the range $2–20\,\mathrm{gcm^{-3}}$. The closer the value of the density of a fiber to its theoretical value, the smaller will be the amount of microvoids in the fiber and the closer it will be to its theoretical composition. For example, the theoretical bulk density of carbon is $2.1\,\mathrm{g\,cm^{-3}}$, while the actual bulk density of carbon fiber can range from about $1.6–1.8\,\mathrm{g\,cm^{-3}}$ to close to $2\,\mathrm{g\,cm^{-3}}$ for highly aligned graphitic fibers obtained from a pitch precursor. The reason for this is that the pitch-based graphitic fibers have fewer noncarbon atoms and microvoids than other carbon fibers.

10. Mechanical and Thermal Properties

A summary of the mechanical characteristics of important fibers is given in Table 1. It should be pointed out that these values refer to the axial direction of the fiber. As mentioned above, most fibers have characteristics that are heavily biased along their axis, i.e., they are highly anisotropic. In some cases (carbon fibers in particular) they are stronger and stiffer along their axial direction than in the transverse or radial direction. Note also that the properties in compression generally differ from those in tension. Some fibers conduct: obviously copper wires, but also some carbon fibers. Carbon fibers also have a negative coefficient of thermal expansion.

11. Conclusions

Materials in a fibrous form have been used for a long time. Fiber yarns have been used for making fabrics, ropes, cords, etc., since prehistoric times, long before people had any idea of the internal structure of these materials. By any measure, fiber-related industrial activity worldwide represents a very important sector of the world's economy. Fibrous materials are, in one form or another, part of our daily life. In particular, the availability of strong and stiff fibers has resulted in very important developments in the field of fiber-reinforced composites, spanning the fields of sporting equipment, marine and civil construction, the aerospace industry, and defense applications involving aircrafts, rocket nozzles, and nose cones of missiles and space vehicles.

Bibliography

Chawla K K 1993 *Ceramic Matrix Composites*. Chapman & Hall, London

Chawla K K 1998a *Fibrous Materials*. Cambridge University Press, Cambridge

Chawla K K 1998b *Composite Materials: Science & Engineering*, 2nd edn. Springer, New York

Chawla K K, Xu Z R, Hlinak A, Chung Y W 1993. In: *Advances in Ceramic Matrix Composites*. American Ceramic Society, Westerville, OH, p. 725

Chou T-W, Ko F K (eds.) 1989 *Textile Structural Composites*. Elsevier, Amsterdam

Clyne T W, Withers P J 1993 *An Introduction to Metal Matrix Composites*. Cambridge University Press, Cambridge, UK

DeBolt H E, Krukonis V J, Wawner F E 1974. In: *Silicon Carbide 1973*. University of South Carolina Press, Columbia, SC, p. 168

Dresher W H 1969 *J. Metals* **21**, 17

Hull D, Clyne T W 1996 *An Introduction to Composite Materials*, 2nd edn. Cambridge University Press, Cambridge, UK

Perez G, Ziabicki A, Kawai A (eds.) 1985 *High Speed Fiber Spinning*. Wiley, New York, p. 333

Piggott M R 1980 *Load Bearing Fibre Reinforced Composites*. Pergamon, New York

Suresh S, Needleman A, Mortensen A M (eds.) 1993 *Fundamentals of Metal Matrix Composites*. Butterworth-Heinemann, Boston

Taya M, Arsenault R J 1989 *Metal Matrix Composites*. Pergamon, New York

Vassilatos G, Knox B H, Frankfort H R E 1985. In: Ziabicki A, Kawai A,(eds.)*High Speed Fiber Spinning*. Wiley, New York, p. 367

Wax S G 1985 Ceramics research at DARPA. *Bull. Am. Ceram. Soc*. **64**, 1096

Yajima S 1980 Development of ceramics, especially silicon carbide fibres from organosilicon polymers. *Phil. Trans. R. Soc. London* **A294**, 419

Yajima S, Okamura K, Hayashi J, Omori M 1976 Synthesis of continuous silicon carbide fibres. *J. Am. Ceram. Soc* **59**, 324

K. K. Chawla

University of Alabama at Birmingham, Birmingham, Alabama,USA

Foams, Microrheology of

Foams are low-density cellular materials, formed by introducing gas bubbles into liquids through physical or chemical means. They are encountered daily in food, beverage, personal care, and household items, and are used in firefighting, mineral processing, and petroleum recovery, all of which illustrate the broad range of materials and applications that motivates the development of foam science and technology (Bikerman 1973, Kraynik 1988, Prud'homme and Khan 1996, Weaire 1996, Gibson and Ashby 1997, Sadoc and Rivier 1999, Weaire and Hutzler 2000, Perkowitz 2000). Foams are prototypical multiphase materials; their structure involves multiple length scales and their behavior embodies many physical mechanisms. A micromechanical point of view provides a theoretical foundation for establishing structure–property–processing relationships to guide the use and development of foams. Cellular solids that begin life as gas–liquid foams retain much of the structure that evolves in the liquid state. Polyhedra packed to fill space underlie the geometry of all low-density foams.

This article focuses on the structure and rheology of soap foams under quasistatic conditions. Dry soap foams consist of polyhedral gas bubbles separated by thin liquid films. They can be modeled as networks of minimal surfaces. Even though these foams are composed primarily of two Newtonian fluids—water and air—they behave like soft solids because the cells are jammed together. Elastic response occurs when the foam is strained and the network of thin films deforms without changing topology; strain energy is stored as increased surface area. Plastic response occurs when the film network loses stability as it deforms, which provokes irreversible cascades of topological transitions that release energy. The foam stress and energy density scale as $\sigma/V^{1/3}$, where σ is surface tension and V is average cell volume. Foams with smaller cells are stiffer and stronger as measured by the shear modulus, G, and yield stress, σ_y, respectively (Saint-Jalmes and Durian 1999). Foams are far from equilibrium: films rupture, gravity encourages liquid drainage (Verbist *et al.* 1996), and gas diffuses between cells causing coarsening.

1. *Foam Structure*

Under static conditions, the local geometry of dry soap foams satisfies Plateau's laws: (i) each face on a polyhedron is a minimal surface with uniform mean curvature that balances the pressure difference between neighboring cells; (ii) three faces meet at each cell edge, a degenerate Plateau border, at equal dihedral angles of 120°; (iii) four edges meet at each cell vertex at equal tetrahedral angles of $\cos^{-1}(-1/3) = 109.47°$. These conditions imply that no face can be a flat polygon with straight edges because no such polygon has all vertex angles equal to the tetrahedral angle. Lacking analytical solutions for the shape of the minimal surfaces, the Surface Evolver (Brakke 1992) has become the standard computer software for calculating foam structures (it is available from: http://www.susqu.edu/facstaff/b/brakke/evolver/). All foams discussed here are modeled as spatially periodic structures with N cells in a representative volume or "unit cell."

The Kelvin cell shown in Fig. 1 is a natural starting point for modeling three-dimensional foams because it is the only polyhedron known to form a perfectly ordered structure that satisfies Plateau's laws. It sits on a b.c.c. lattice and has 14 faces: 6 flat quadrilaterals with curved edges and 8 nonplanar hexagons. The Williams foam contains two 14-hedrons with the same shape but different orientations. These cells arrange into alternating layers and have two quadrilateral, eight pentagonal, and four hexagonal faces. The Weaire–Phelan foam uses less surface to divide space into equal-volume cells than any known structure. It contains two pentagonal dodecahedra and six 14-hedrons, and belongs to a class of structures known as tetrahedrally close packed (t.c.p.) to crystallographers and Frank–Kasper to materials scientists (Rivier 1994). All t.c.p. foams are composed of only four types of cells with 12, 14, 15, or 16 faces. Each cell has 12 pentagons; the other faces are hexagons that do not touch each other. The three t.c.p. foams with cubic symmetry are shown in Fig. 1; they and their alternative names are Weaire–Phelan (A15), Friauf–Laves (C15), and Bergman (T).

Ordered foams are useful models of three-dimensional structure but real foams are disordered. Matzke (1946)carefully prepared and examined monodisperse foams (equal-volume cells) and classified the different polyhedra according to face content. Data on 600 cells from the foam interior are summarized in Fig. 2. The first step in modeling random foams involves molecular dynamics simulation to produce a random close packing of N hard spheres with adjustable size distribution. The center and radius of each sphere are used to construct weighted Voronoi polyhedra whose size distribution reflects that of the spheres. The Surface Evolver is then used to relax the Voronoi structures under the constraint of fixed cell volumes. This process involves many topological transitions before structures that satisfy Plateau's laws are achieved. Random, monodisperse foams are produced by starting with equal-size spheres and using equal cell volumes as constraints. Typical distributions of polyhedra with F faces and polygons with n edges are shown in Fig. 2. Matzke (1946) found averages of $\bar{F} = 13.70$ and $\bar{n} = 5.124$. Ten simulations with $N = 1000$ produced excellent agreement with experiment: $\bar{F} = 13.86$ and $\bar{n} = 5.134$. Triangular faces are common among the Voronoi cells but very rare in fully relaxed monodisperse foams. Matzke did not find any triangular faces or a single Kelvin cell. The simulations produced a few Kelvin cells but they are very rare (0.2%), which is not surprising since pentagonal faces are so common. Figure 2 shows some of the more common cells. The

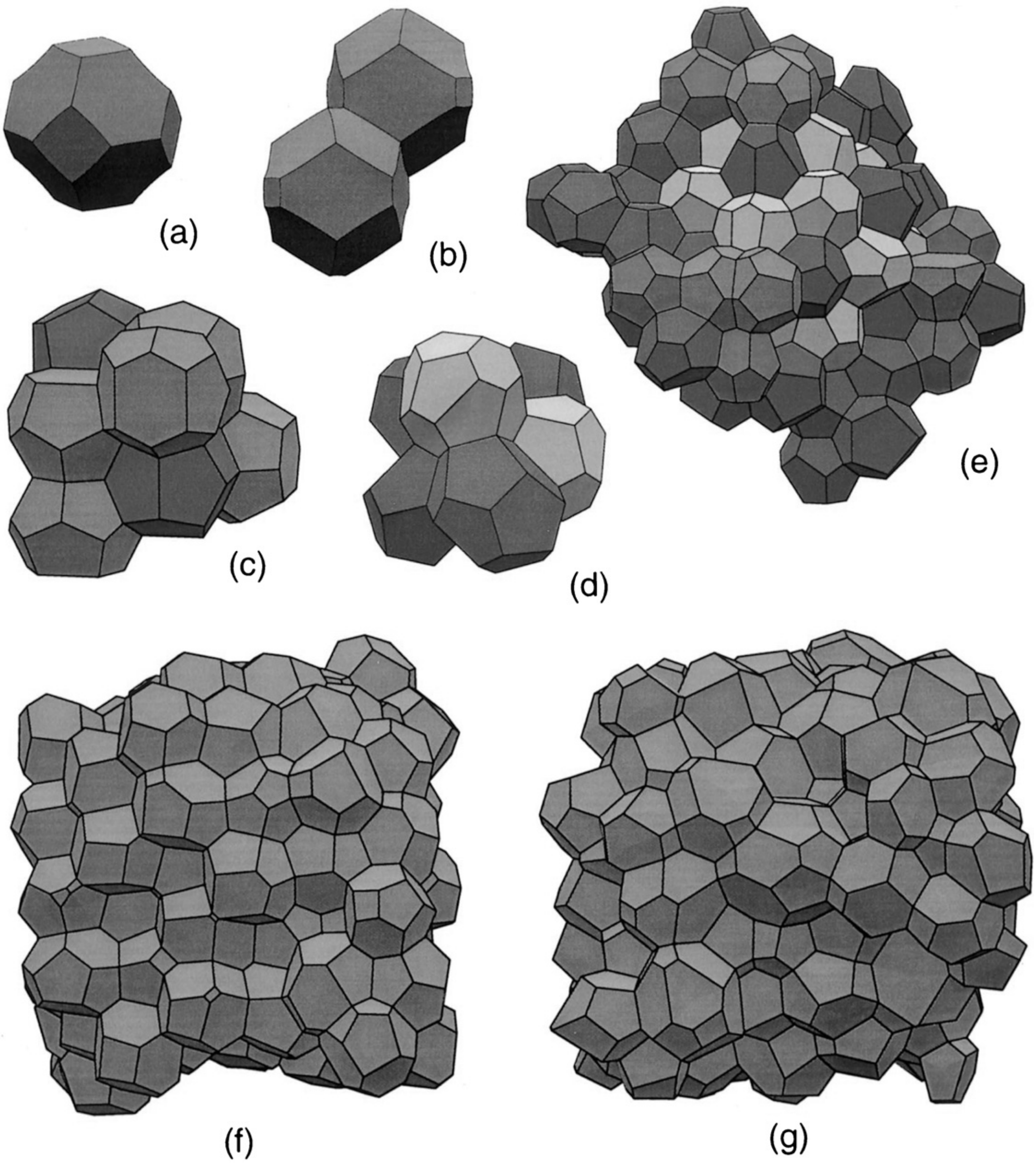


Figure 1
Foam structures: (a) Kelvin cell, (b) Williams, (c) Weaire–Phelan, (d) Friauf-Laves, (e) Bergman, (f) random monodisperse, (g) random polydisperse.

predominant polyhedron recorded by Matzke (19.7%) was a 13-hedron (1, 10, 2); the notation (n_4, n_5, n_6) indicates the number of n-gons. There is only one way to assemble this cell: the two hexagons are attached to opposite sides of the quadrilateral. Among the more common cells in the simulations (13.2%) were 14-hedrons (2, 8, 4) that include the Williams cell but the faces can be combined in 11 different ways. Pentagonal dodecahedra, which are often considered to be the prototypical foam cell, are well represented but not the most common. Figure 2 also contains F and n distributions for two polydisperse foams whose cell volumes vary by over a factor of 10. They contain a broader distribution of

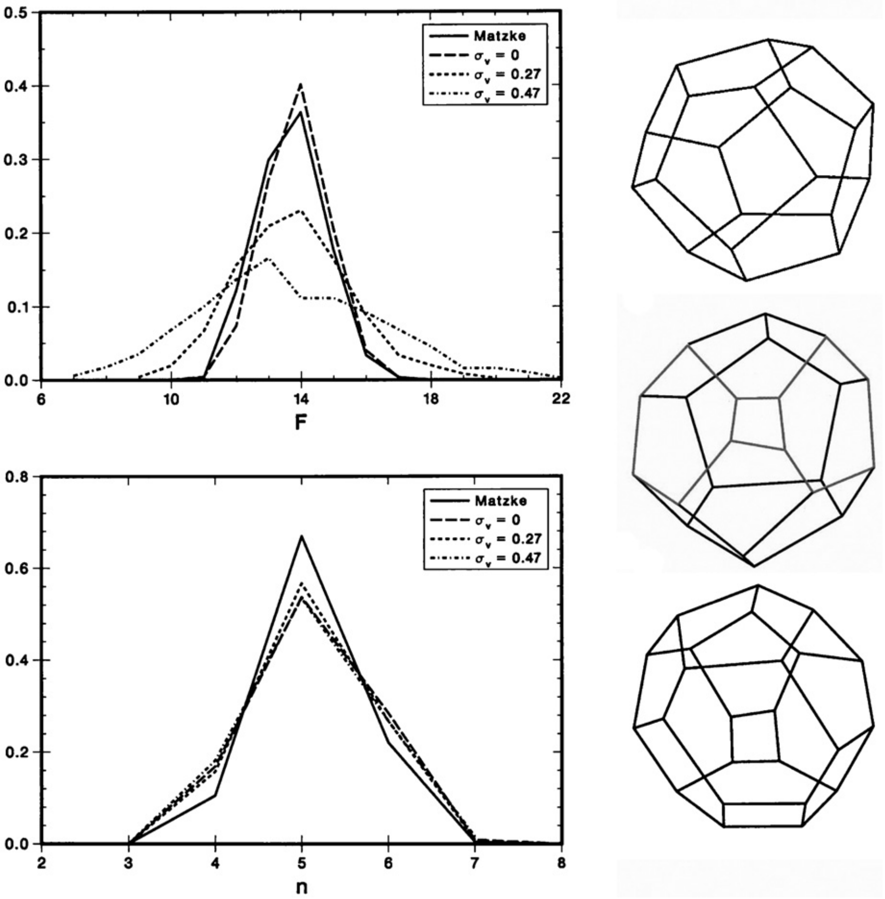


Figure 2
Distributions of polyhedra with F faces and polygons with n edges in random foams. Some common cells in random monodisperse foam: pentagonal dodecahedron, 13-hedron (1, 10, 2), and 14-hedron (2, 8, 4).

F-hedrons than monodisperse foams; smaller cells tend to have fewer faces and vice versa. The distribution of n-gons is surprisingly similar to monodisperse foams. Most of the faces have four to six edges when the cell-size distribution is not too broad. A large number of extremely small tetrahedral cells would significantly increase triangular faces.

2. Rheology

The Surface Evolver can be used to deform spatially periodic foams, compute the minimal structures, and evaluate the stress. The *effective isotropic* shear modulus of ordered foams ranges from $G = 0.78\sigma/V^{1/3}$ for the Kelvin foam to $G = 0.87$ for t.c.p. foams. Calculations for random monodisperse foams with $N \leqslant 216$ give $G = 0.80 \pm 0.02$; polydisperse foams are slightly softer. Experimental studies on wet foams and concentrated liquid–liquid emulsions provide consistent estimates of the shear modulus in the dry limit: $G = 0.82$, which is in excellent agreement with simulations. No dependence on polydispersity has been reported from measurements. Coincidentally, the shear moduli of Kelvin and random foams are very close. By contrast, a calculation based on affine deformation predicts $G > 1.4$

Elastic–plastic behavior occurs when a dry foam is subjected to quasistatic simple shearing flow. Figure 3 shows the evolution of shear stress and surface energy density with strain for a random monodisperse foam with 64 cells. The fluctuations correspond to storage and release of strain energy. The jumps occur when cell edges go to zero length, which violates Plateau's laws and triggers cascades of local topological transitions called *T1*s. The irreversible sequence of events that defines a *T1* cascade and dissipates energy is a fundamental mechanism for plasticity in jammed

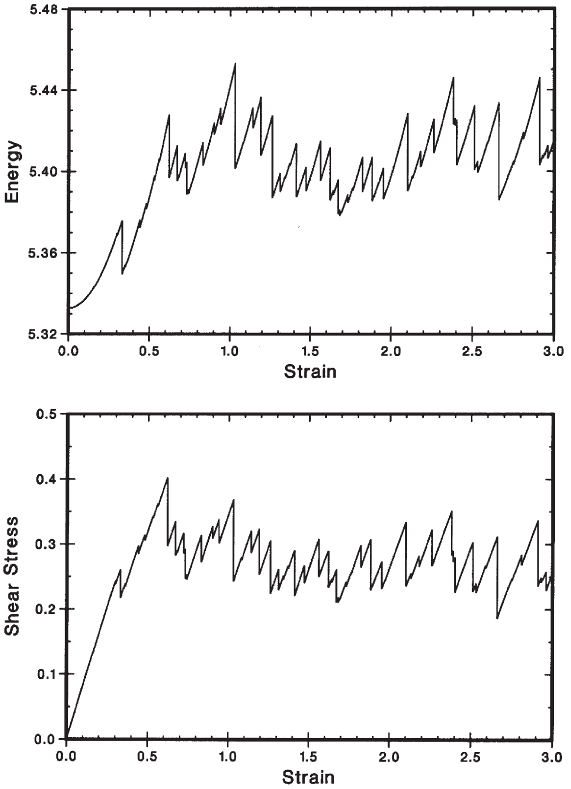


Figure 3
Evolution of shear stress and surface energy density for a random foam in simple shearing flow.

fluid–fluid systems. The stress saturation at large strains determines the viscometric functions of a dry foam in the limit of zero shear rate. The shear stress "plateau", in particular, determines a yield stress σ_y, which is different from the elastic limit defined by the first *T1* avalanche. Measurements of yield stress vary significantly, $\sigma_y = 0.11$–0.22. Simulations suggest a slightly higher value: $\sigma_y \approx 0.3$. Considering the uncertainties that surround the measurement of nonlinear rheology in these relatively fragile fluids, the agreement is reasonable. Stress–strain curves for Kelvin and Weaire–Phelan foams exhibit much larger fluctuations and strong dependence on foam orientation (Reinelt and Kraynik 2000). Nonlinear properties obtained from simulations on small systems are unreliable.

3. *Diffusive Coarsening*

Gas diffusion between cells causes foam coarsening, a prevalent destabilizing mechanism related to Ostwald ripening. A scaling argument suggests that R, the average cell "radius," grows with the square root of

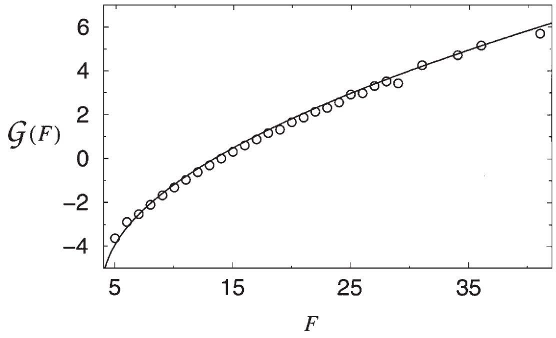


Figure 4
Average growth rate of F-hedrons in a polydisperse foam compared against the mean-field theory (solid line) for $\mathcal{G}(F)$ discussed in the text.

time in two or three dimensions; therefore, diffusive coarsening causes a foam to soften with time since stress scales as σ/R. In two dimensions, the growth rate of a polygonal cell with n sides is given by von Neumann's law:

$$\frac{\mathrm{d}A}{\mathrm{d}t} = D(n-6) \tag{1}$$

where A is the cell area and D is an effective diffusion coefficient. This exact relationship shows that the growth rate is independent of cell size and solely dependent on n. Cells with six sides, the average in two dimensions, neither grow nor shrink. The situation in three dimensions is not as simple. The growth rate of individual cells does not depend on F alone, so efforts have focused on finding an *average* growth law. Dimensional analysis and physical arguments suggest

$$V_F^{-1/3}\frac{\mathrm{d}V_F}{\mathrm{d}t} = D\mathcal{G}(F) \tag{2}$$

where V_F is the average volume of cells with F faces and $\mathcal{G}(F)$ is some dimensionless function of bubble shape that depends on F and other features of the foam structure. The instantaneous growth rate of individual cells can be evaluated from Surface Evolver calculations, which provide cell pressures and geometry. Figure 4 shows average growth rate plotted against F for a polydisperse foam with 512 cells. The data are well represented by a mean-field theory for which $\mathcal{G}(F \rightarrow \infty) \sim \sqrt{F}$ (Hilgenfeldt *et al.* 2000).

4. *Wet Foams*

Unlike the dry limit, a perfectly ordered wet foam can have different structure besides a wet Kelvin cell, e.g., the wet rhombic dodecahedron associated with f.c.c. packing (see Fig. 5). This indicates that Plateau's laws do not pertain to wet foams with finite liquid volume

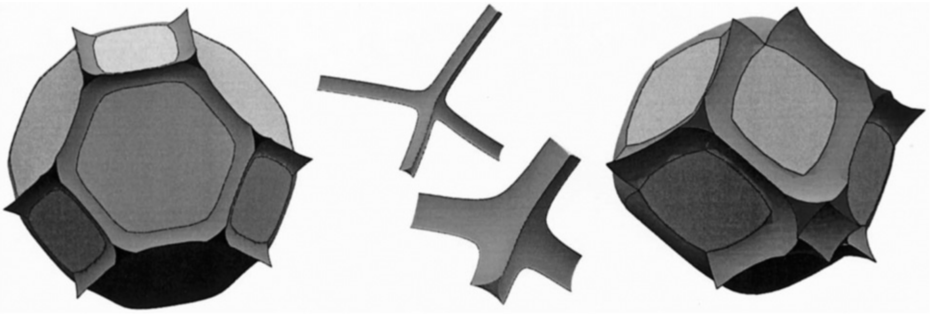


Figure 5
Wet Kelvin cell, $\phi = 0.01$; Plateau border segments from a Kelvin cell, $\phi = 0.001$, 0.01; wet rhombic dodecahedron, $\phi = 0.06$.

fraction ϕ. The connectivity of Plateau borders that form along cell edges can exceed four, permitting a richer variety of foam topologies. Wet Kelvin cells on a b.c.c. lattice can be stretched to form wet rhombic dodecahedra on an f.c.c. lattice (Kraynik *et al.* 1999). The topological transition between the two types of cells also occurs during simple shearing flow. Surface Evolver models of perfectly ordered wet foams capture some qualitative features of real systems, such as the shear modulus and yield stress decreasing with ϕ. However, accurate calculations for wet foams with random structure and large N are beyond current computing resources.

5. *Summary*

This brief survey of static foam structure and rheology from a micromechanics point of view reveals a firm theoretical foundation for developing structure–property–processing relationships for foamed materials. The interplay between viscosity and surface tension and its influence on foam microrheology exemplifies further challenge and opportunity in modeling foam processing. The fluid mechanics of bubbles growing and deforming in a viscous fluid provides a link to the micromechanics of cellular solids by setting microstructure.

See also: Elastic Behavior of Cellular Steels*

Bibliography

Bikerman J J 1973 *Foams*. Springer-Verlag, New York

Brakke K A 1992 The surface evolver. *Exp. Math.* **1**, 141–65

Gibson L J, Ashby M F 1997 *Cellular Solids: Structure and Properties*. Cambridge University Press, Cambridge, UK, 2nd edn.

Hilgenfeldt S, Kraynik A M, Koehler S A, Stone H A 2001 An accurate von Neumann's law for three-dimensional foams. *Phys. Rev. Lett.* **86**, 2685–88

Kraynik A M 1988 Foam flows. *Annu. Rev. Fluid Mech.* **20**, 325–57

Kraynik A M, Neilsen M K, Reinelt D A, Warren W E 1999 Foam micromechanics: structure and rheology of foams, emulsions, and cellular solids. In: Sadoc J F, Rivier N (eds.) *Proc. School on Foams, Emulsions, and Cellular Materials*. Kluwer, Dordrecht, The Netherlands, pp. 259–86

Matzke E B 1946 The three-dimensional shape of bubbles in foam—an analysis of the role of surface forces in three-dimensional cell shape determination. *Am. J. Botany* **33**, 58–80

Perkowitz S 2000 *Universal Foam: From Cappuccino to the Cosmos*. Walker, New York

Prud'homme R K, Khan S A (eds.) 1996 *Foams: Theory, Measurements, and Applications, Surfactant Science Series*. Marcel Dekker, New York, Vol. 57

Reinelt D A, Kraynik A M 2000 Simple shearing flow of dry soap foams with tetrahedrally close-packed structure. *J. Rheology* **44**, 453–71

Rivier N 1994 Kelvin's conjecture on minimal froths and the counter-example of Weaire and Phelan. *Phil. Mag. Lett.* **69**, 297–303

Sadoc J F, Rivier N (eds.) 1999 Foams and emulsions. In: *Proc. School on Foams, Emulsions, and Cellular Materials*. Kluwer, Dordrecht, The Netherlands

Saint-Jalmes A, Durian D J 1999 Vanishing elasticity for wet foams: equivalence with emulsions and role of polydispersity. *J. Rheology* **43**, 1411–22

Verbist G, Weaire D, Kraynik A M 1996 The foam drainage equation. *J. Phys.: Condens. Matter* **8**, 3715–31

Weaire D (eds.) 1996 *The Kelvin Problem: Foam Structures of Minimal Surface Area*. Taylor and Francis, London

Weaire D, Hutzler S 2000 *The Physics of Foams*. Oxford University Press, Oxford, UK

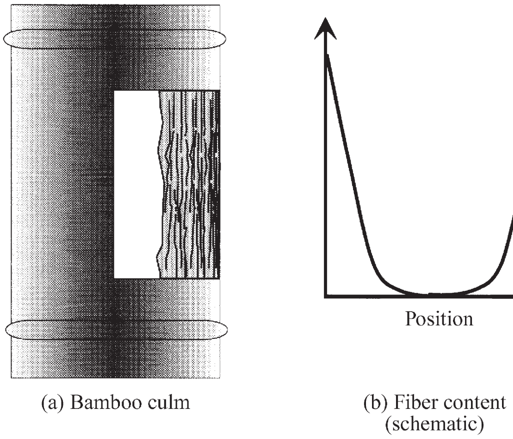
A. M. Kraynik[a] and D. A. Reinelt[b]
[a] *Sandia National Laboratories, Albuquerque, New Mexico, USA*
[b] *Southern Methodist University, Dallas, Texas, USA*

Functionally Graded Materials

A functionally graded material (FGM, or sometimes also "gradient material") is characterized by a gradual change of material properties with position. The property gradient in the material is caused by a position-dependent chemical composition, microstructure, or atomic order. The spatial extension of the gradient may differ: in a bulk FGM the property variation extends over a large part of the material, whereas in a graded coating or joint it is restricted to the surface of the material or a small interfacial region.

Although FGMs attracted scientific interest only towards the end of the twentieth century, these materials are not new. In fact, spatial variations in the microstructure of materials have been exploited for millions of years by living organisms. In many structures found in plants, microstructural gradients are formed in order to produce optimum structural and functional performance with minimum material use. An example is the culm of bamboo, which consists of high-strength natural fibers embedded in a matrix of ordinary cells (Fig. 1(a)). The fiber content is not homogeneous over the entire cross-section of the culm but decreases from outside to inside (Fig. 1(b)). This gradation in fiber content is a natural adaptation of the plant to flexural loads—the fiber content is high only in those sections where the highest stresses occur.

There are also some examples of early synthetic materials taking advantage of property gradients such as case-hardened steels in which a hard surface is combined with a tough interior. Nevertheless, it was only in 1987 that Japanese scientists conceived graded materials as a new concept and defined an FGM as an "inhomogeneous composite, in which the material's mechanical, physical, and chemical properties change continuously, and which have no discontinuities within the material" (Hirai *et al.* 1987).



(a) Bamboo culm

(b) Fiber content (schematic)

Figure 1
Cross-section of a bamboo culm: (a) cross-section; (b) distribution of fibers across the culm (after Amada *et al.* 1995).

Meanwhile it has become common practice to use the term FGM not only for composites, but also for all materials with a macroscopic property gradient (see definition at the beginning of the article). Thus, materials containing only one phase such as graded glasses or graded single crystals are also FGMs. In such materials the spatial property variation is caused by a gradient in chemical composition. Nevertheless, the vast majority of FGMs are composite materials with a macroscopic microstructural gradient. For example, the composite may contain a spatially varying volume fraction of one of the phases (Fig. 2(a)). In this case, the gradient material can be conveniently described by the use of a transition function $f(x,y,z)$, where f is the volume fraction of one of the phases as a function of position. In many practical cases the compositional variation will be restricted to one coordinate, z, and the different gradients can then be described by a so-called transition function of the type:

$$f(z) = \left(\frac{z}{d}\right)^p \tag{1}$$

where f denotes the volume fraction of one of the phases, d is the thickness of the graded region, and p is the so-called gradation exponent. However, a composition gradient is not inherent to all FGMs. Microstructural gradients may also be obtained in composites by changing the shape (Fig. 2(b)), orientation (Fig. 2(c)), or size (Fig. 2(d)) of the dispersed phase.

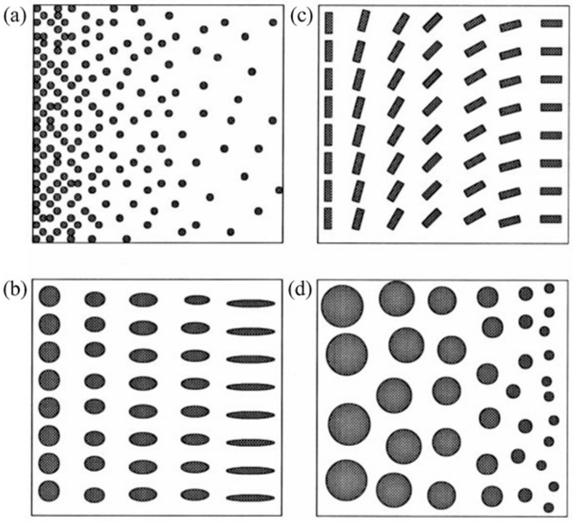


Figure 2
Different types of functionally graded composites. Gradient of: (a) fraction, (b) shape, (c) orientation, and (d) size of material.

1. Functions of Property Gradients in Materials

Properly designed property gradients in materials lead to a performance that cannot be achieved with a homogeneous material or by the joining of two different materials. The gradient can thus be regarded as an additional material design parameter which may be adapted and optimized to meet the requirements of a particular application. This leads to the unusual situation that the performance of the material is strongly influenced by geometric factors such as the shape of the material, and the extension and profile of the property gradient. Thus, a gradient material cannot be fully characterized by materials constants only, and already comprises certain aspects of the design of a component.

Most of the theoretical research on FGMs has been devoted to their macroscopic mechanical and thermal behavior and there has been considerable progress in understanding the effects of gradients on the stress distribution within such materials (reviewed by Suresh and Mortensen 1998). A gradation of the elastic modulus and/or of the thermal expansion coefficient changes the stress distribution in a component when it is mechanically or thermally loaded. A simple but illustrative example is a trilayered beam of infinite length in the x-direction, and unit width in the z-direction. In the y-direction it consists of two homogeneous layers of thickness h_1 and h_2 and a graded interlayer of thickness g (Fig. 3). For this simple geometry the stresses inside the beam can be conveniently calculated using classical continuum mechanics. It is assumed that plane stress conditions prevail in the center of the beam. If the beam is subjected to a uniform temperature change from a stress-free temperature, a strain gradient in the thickness direction y will evolve due to the difference in thermal expansion of the two materials and the beam will bend. The only nonzero stress in the center of the beam is $\sigma_{xx}(y)$. The stress distribution $\sigma_{xx}(y)$ for a Cu/Al_2O_3 beam of 5mm thickness after a uniform temperature change of 100 K is shown in Fig. 4. In the case of a simple bimaterial ($g=0$, $h_1=h_2=2.5$mm) a large stress jump with high compressive stresses in the Al_2O_3 and high tensile stresses in the copper is predicted near the interface. If a layer that possesses a linear variation of composition from copper to Al_2O_3 is introduced in the central part of the beam ($h_1=h_2=$1mm, $g=3$ mm, exponent p of the transition function $=1$), the stresses $\sigma_{xx}(y)$ will be substantially reduced and the abrupt change in sign of the stress disappears. Instead, a gradual transition from tensile to compressive stress is observed and the stress becomes zero in the center of the beam. If desired, the exact location y_0 where $\sigma_{xx}(y)$ vanishes can be shifted to a different position by choosing an appropriate transition function. If the linear gradation of the thermal expansion coefficient extends over the whole plate thickness ($h_1=h_2=0$, $g=5$ mm), the thermal stresses will even vanish completely. However, a graded layer that is thin compared to the homogeneous layers ($g \ll h_1+h_2$) cannot affect the force balance inside the whole beam and the stress reduction in the nongraded part will be negligible.

It should be pointed out that a stress reduction in bimaterials is not a general feature of FGMs. For

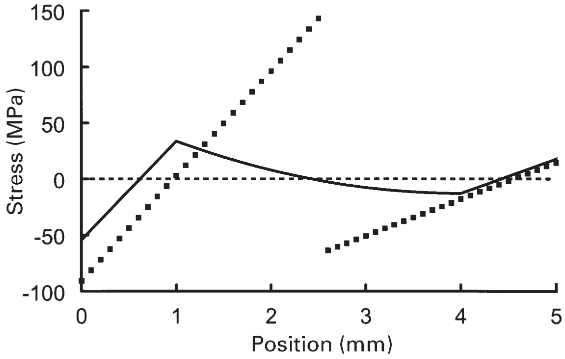


Figure 4
Stresses σ_{xx} along the y-axis of a thin Cu/Al_2O_3 bar after a temperature change of 100K from the stress-free temperature: ↑, direct joint; —, with graded layer of 3mm thickness in the center; - - -, linear gradation across the entire thickness of the bar.

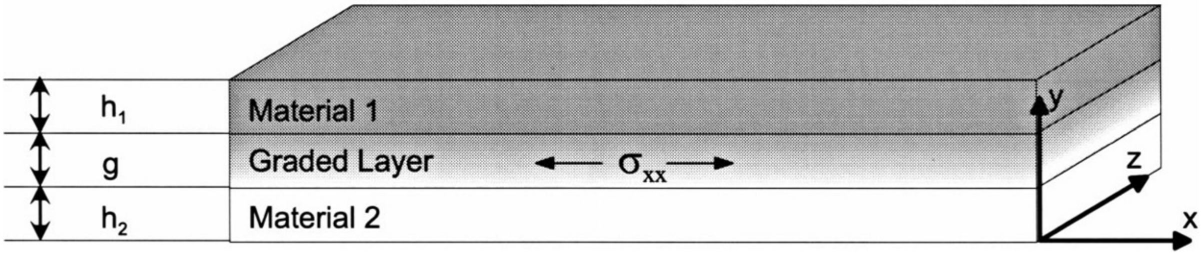


Figure 3
Geometry of FGM bar used for stress calculation.

arbitrary geometry and loading situation, the gradient will not always lead to a decrease in stress. However, a properly chosen gradient in thermoelastic properties can aid in redistributing stresses in a favorable manner. This can be an important feature in situations where the size or shape of a component is restricted.

So far, only stresses in the interior of an FGM have been considered. There is another important effect of the gradient on the stresses at the free edge of the beam, in particular at the locations where the interface intersects the free boundaries. Assuming elastic material behavior, the stresses σ become singular (infinite) at these locations according to:

$$\sigma_{ij}(r,\Theta) = \frac{K}{r^{\lambda}} f_{ij}(\Theta) \tag{2}$$

where K is a constant, $f(\Theta)$ is a function describing the angular distribution, r is the distance from the free edge, and λ is the stress exponent. A similar stress singularity for the stress σ_y and the shear stress τ_{xy} is also predicted at an elliptical pore located at the interface of dissimilar materials (Yang *et al.* 1997). The stress concentrations at the free edges or pores are sites where the nucleation of critical cracks (leading to delamination) is likely to occur. If a graded zone with continuously varying mechanical properties is introduced, the stresses at the free edge of the interface become finite (Yang 1998). It was also demonstrated that the introduction of a gradient leads to finite values of σ_y and zero shear stresses τ_{xy} at a pore located at an interface. It is therefore expected that delamination cracking in bimaterial joints or coatings can be substantially reduced or suppressed by the introduction of a property gradient. Indeed, delamination of a TiC/Ti coating on a silicon substrate was shown to be suppressed by a properly chosen gradient (Dahan *et al.* 1999).

In the above, only the effect of a gradient in elastic and thermal properties on the stress distribution in a crack-free material was discussed. However, a gradient will not only affect crack initiation, but also the mechanical energy release rate of an existing crack. Fracture mechanical calculations show that an edge crack propagating along the interface of a homogeneous coating possesses a higher mechanical energy release rate compared to a crack in a graded coating (Erdogan 1995). Furthermore, a suitable local variation of fracture toughness and subcritical crack growth behavior can improve the fatigue life of a component considerably. It was demonstrated that a gradual transition from a fully aged, brittle, high-strength titanium alloy at the surface to a ductile, low-strength titanium alloy is produced by roller burnishing and a subsequent heat treatment. The graded structure retained by the surface treatment improves considerably the high-cycle fatigue strength of the titanium alloy (Berg and Wagner 1999).

The advantages of gradients are not restricted to structural applications. Gradient materials can display functional properties that cannot be achieved by a homogeneous material or directly joined material. Indeed, one of the earliest uses of property gradients in functional materials were graded refractive index lenses and graded refractive index glass fibers for data transmission which allow outstanding transfer rates even for long transmission paths (Uchida *et al.* 1970). Additionally, the absence of sharp interfaces in FGMs offers the chance to eliminate undesired effects of interfaces such as reflections of electromagnetic or acoustic waves or pile-up of dislocations in heteroepitaxy. Such tailor-made properties make FGMs attractive for many acoustic, optic, electronic, magnetic, piezoelectric, thermoelectric, and photovoltaic applications (see Sect. *3*).

2. Processing of FGMs

A large part of the research work on FGMs has been dedicated to processing and a large number of production methods have been developed (reviewed by Suresh and Mortensen 1998, Neubrand and Rödel 1997). Most of the processes for FGM production are based on a variation of conventional processing methods which are already well established. Methods that are capable of accommodating a gradation step include powder processing, thermal spraying, melt processing, and coating processes. Which of these production methods is most suitable depends mainly on the material combination, type of transition function required, and geometry of the desired component (Table 1).

In powder processing, the gradient is generally produced by mixing different powders in variable ratios and stacking the powder mixtures in separate layers. The thickness of the separate layers is typically between 0.2 mm and 1 mm. However, even thinner layers with a quasicontinuous composition can be obtained by employing centrifugal powder metallurgy (Delfosse and Ilschner 1992). In this process the powder mixtures and a liquid binder are projected by centrifugal forces onto a support ring, where a solid green body is produced. Quasi-continuous gradients may also be formed by spraying powder suspensions or slip casting having a variable slip composition. The major problem in powder processing of FGMs is the sintering step, as the sintering rates of the different components may differ by orders of magnitude leading to cracking or warping of the compacts. Numerous efforts to counterbalance unequal sintering activity have been undertaken, including adaptation of particle size of the powder components in such a way that the sintering rate becomes similar over a wide range of temperature, and addition of a sintering aid to the more refractory component to increase its sintering rate. Local sintering rates can also be

Table 1
Characteristics of processing methods for FGMs.[a]

	Powder metallurgy	Thermal spraying	Coating processes	Melt processing
Processing of bulk FGMs	+	O	–	+
Processing of graded joints	O	–	–	O
Processing of graded coatings	–	+	+	–
Variability of transition function	+	+	+	O
Achievable density	O	–	O	+
Processing of complex shapes	O	–	–	+
Example	ZrO_2/stainless steel bulk FGM	YSZ/NiCrAlY thermal barrier coating	TiC/TiN wear-resistant coating	W/Cu heat sink material

[a] +: good, very suitable; O: fair, suitable; –: poor, unsuitable.

equalized by sintering in a temperature gradient which can be produced by laser surface heating or plasma-activated sintering in a stepped or tapered die. Hot pressing is also an efficient means to prevent deformation.

Thermal spraying has been frequently used to produce FGM coatings. Thermal spraying of FGMs offers the possibility to combine highly refractory phases with low-melting metals and allows for direct setting of the gradation profile. The powders for each component can be delivered through multiple feeders using a single torch or using separate torches. Materials produced by thermal spraying have a porosity between 5% and 20%. It is thus particularly attractive for applications where this porosity can be tolerated or is even advantageous such as in graded NiCrAlY/YSZ thermal barrier coatings. Thinner graded coatings have been produced by vapor deposition processes such as conventional physical vapor deposition (PVD), electron beam physical vapor deposition (EBPVD), magnetron sputtering, chemical vapor deposition (CVD), and galvanic plating. In all of these processes the composition gradient is produced by a controlled release of the components from a source such as an anode, ingot, liquid, or gas reservoir.

Bulk FGMs containing a low-melting metal can be favorably produced by melt processing. Particles of the more refractory phase are added to the melt and the gradient is produced by sedimentation or centrifuging during solidification. The more refractory phase can also be provided as a graded preform which is infiltrated with the melt. Processes used to prepare graded preforms include sequential slip casting of dispersions with a variable content of short fibers, weaving long-fiber preforms with a variable content of spacer particles, electrochemical generation of porosity gradients in homogeneous preforms, and inhomogeneous compression of polymer foams.

3. *Potential Applications*

Due to the development of the FGM concept only in the 1990s, the practical use of FGMs is still very limited. There are, however, a few examples of early applications of materials that are FGMs by definition. Already in the 1960s glass/metal joints with a gradient in thermal expansion coefficient were used to reduce thermal stresses at the interface of the two materials and thus prevent failure of the joint. Glasses with graded refractive index were developed in the early 1970s and were probably the first materials in which a very specific and precise gradation profile was essential for application. A circular plate with a gradient in refractive index can act as a miniature focusing lens with no spherical aberration. Glass rods with a radial index gradient drawn to a fiber can be used in wide-band optical fiber cables. Another well-established application of FGMs is in polymer skin foams with a dense surface layer and an increasing porosity towards the interior. These foams can provide high impact strength at low weight in applications such as instrument panels.

All the materials mentioned above are graded glasses or polymers, which are comparatively easy to manufacture. Gradient composite materials consisting of components with strongly different properties, i.e., graded ceramic/metal or ceramic/ceramic composites, have been manufactured only since the 1990s and tests of graded components from such materials in a real environment are still a rarity. Therefore, they are not widely used in commercial products. Nevertheless, they are expected to provide useful solutions for challenging engineering problems (Miyamoto *et al.* 1999). A few promising examples of FGMs for structural applications are given below. Table 2 gives examples of some proposed functional applications.

Table 2
Proposed applications of FGMs as functional materials.

Proposed application	Proposed FGM	Function of gradient
Energy conversion		
Thermoelectric converter	$Bi_2Te_3/PbTe/SiGe$	Improved conversion efficiency
Solid oxide fuel cell	$La_{0.85}Sr_{0.15}MnO_3/La_{0.5}Sr_{0.5}CoO_3$	Increased power output
Radiation emitter	AlN/W	Improved emissivity
Acoustics/optics		
Bimorph actuator	$Pb(Ni_{1/3}Nb_{2/3})_{0.5}(Ti_{0.7}Zr_{0.3})_{0.5}O_3/$ $Pb(Ni_{1/3}Nb_{2/3})_{0.7}(Ti_{0.7}Zr_{0.3})_{0.3}O_3$	Reliable joining of piezoelectric/dielectric material
Acoustic coupler	CuMn/WC	Acoustic impedance matching
Optical bandpass filter	SiO_2/TiO_2	Improved transmittance characteristics
X-ray mirror	Si/Ge	Graded lattice constant
Electronics		
Graded emitter heterojunction bipolar transistor	Dopant gradient in Si	Control of bandgap structure
Graded heteroepitaxial structure	Si/SiGe	Reduction of dislocation density
Graded capacitor	$BaTiO_3/SrTiO_3$	Zero-temperature coefficient of capacitance

3.1 Wear Protection

The wear resistance of tools can be substantially increased by a hard surface layer in a compressive stress state (Lin *et al.* 1998). However, if a coating is under compressive stress the risk of delamination is considerable for many material combinations. A WC/Co FGM with a gradual variation in cobalt binder content from the surface to the interior will develop a compressive stress state at the surface. However as explained in Sect. 1 spallation of the top layer is unlikely to occur. Sandvik AB and Sumitomo Electric Industries already use this concept in tools for special applications such as drill bits and throwaway chips. The improved resistance to delamination of an FGM coating was also exploited in an FGM cutting tool where a diamond surface layer was applied to a SiC substrate. Such a tool has a very high cutting precision due to its high stiffness compared to a tool based on a metallic alloy substrate (Li *et al.* 1994).

3.2 Thermal Protection Systems

One of the earliest applications of the FGM concept was the thermal protection of space re-entry vehicles. A SiC/C FGM nose cone for a space plane was developed at the Japanese National Aerospace Laboratory. A nose cone with a SiC protection layer on a C/C composite exposed to a supersonic gas flow at 1900°C deteriorated after the first thermal cycle. In contrast, a nose cone with an additional graded SiC/C interlayer did not show any change after 10 cycles (Sohda *et al.* 1993). Turbine blades of gas turbines must withstand high stresses at elevated temperatures and are thus made from superalloys. The efficiency of the turbine increases with increasing gas inlet temperature and thus the superalloy is coated with a bond coat of NiCrAlY and a heat-insulating layer of ZrO_2 in state-of-the-art gas turbines. The major mechanism limiting the lifetime of such a thermal barrier coating (TBC) is the growth of an oxide layer of alumina at the interface of the bond coat and the ZrO_2 which causes stresses and, ultimately, spallation of the TBC. A gradation of the zirconia coating and the bond coat improved considerably the resistance of the coating to spallation in thermal cycling experiments (Kawasaki and Watanabe 1999). However, it is expected that the oxidation resistance and thus the long-term stability of such a graded system will be insufficient due to the dispersed metal particles in the graded layer. It has therefore been proposed to introduce an additional graded alumina layer between the bond coat and the zirconia. It is not clear yet whether such a design will provide a substantial improvement compared to conventional TBCs. Future use of graded TBCs in turbines is thus uncertain. Likewise, TBCs can be applied to piston crowns and cylinder heads of diesel engines. A 5% reduction in fuel consumption can be achieved by a 2mm thick graded TBC, the lifetime of the graded TBC being much higher than that of an ungraded TBC. Whether graded TBCs will be used in large diesel engines in the future depends presumably on the development of fuel prizes.

3.3 Medical Implants

It is generally acknowledged that a porous surface of a bioactive material such as hydroxyapatite is ideal for rapid bone ingrowth and good mechanical bonding in prosthetic devices. However, for the sake of mechanical stability it is often necessary to use metals as load-carrying materials in implants. As a result, titanium coated with hydroxyapatite is used as an implant material. The adhesion of hydroxyapatite on titanium is only moderate and failure of the interface may occur after implantation. However, it has been demonstrated that the bonding of certain glasses to titanium is quite strong. Therefore, a coating of glass on a Ti–6Al–4V substrate possessing an increasing volume fraction of hydroxyapatite and pores close to its outer surface is an excellent material for cementless artificial joints and dental implants. Such a graded composite showed a good apposition *in vivo*, and the pull-out strength of dog femur implants was three times higher than that of a sandblasted titanium implant (Maruno *et al.* 1999). The pull-out strength increased with time due to bone growth into the interior of the coating layer, and after 76 months of implantation no separation of the glass layer from the titanium was observed. A graded composite of this type therefore has a considerable potential as a cementless artificial joint or endosseous dental implant.

4. Summary

FGMs can be used to avoid problems associated with the presence of an interface in a material: stress singularities due to elastic or thermal property mismatch, poor adhesion, or unwanted reflections at the interface. They can also serve to introduce a designed property profile into a material providing an optimum response to an external field in functional applications. FGMs have been produced successfully with a variety of processing methods including powder, melt and spray processing, as well as various coating techniques. However, for most processing techniques there is a lack of information on the reproducibility of the gradient and the reliability of the produced materials.

This may explain why FGMs are still rare in commercial products despite a considerable number of potential applications that have been proposed in the field of structural and functional materials. An even more important drawback is the additional cost that a gradation process adds to component fabrication. Therefore, in the early years of the twenty-first century it is likely that FGMs will only find applications in components under severe external thermal or mechanical loads where the limits of conventional materials are difficult to advance: thermal protection systems for space vehicles, jet engines or turbines, tools, or medical implants may be candidates. A more widespread application will largely depend on advancements in cost-effective processing of FGMs.

Bibliography

Amada S, Munekata T, Nagase Y, Shimizu N 1995 Hierarchical gradient structures of bamboo. In: Ilschner B, Cherradi N (eds.) *Proc. 3rd Int. Symp. Structural and Functional Gradient Materials*. Presses Polytechniques et Universitaires Romandes, Lausanne, Switzerland, pp. 689–94

Berg A, Wagner L 1999 Near-surface gradient microstructures in metastable beta-titanium alloys for improved fatigue performance. In: Kaysser W A (eds.) *Proc. 5th Int. Symp. Functionally Graded Materials '98*, Materials Science Forum Vols. 308–11. Trans Tech Publications, Switzerland, pp. 307–12

Dahan I, Admon U, Sariel J, Yahav B, Amar M, Frage N, Dariel M P 1999 Functionally graded Ti–TiC multilayers: the effect of a graded profile on adhesion to substrate. In: Kaysser W A (eds.) *Proc. 5th Int. Symp. Functionally Graded Materials '98*, Materials Science Forum Vols. 308–11. Trans Tech Publications, Switzerland, pp. 923–9

Delfosse D, Ilschner B 1992 Pulvermetallurgische Herstellung von Gradientenwerkstoffen. *Materialwiss. Werkstofftech.* **23**, 235-40 (in German)

Erdogan F 1995 Fracture mechanics of functionally graded materials. *Compos. Eng.* **5** (7), 753–70

Hirai T, Sasaki M, Niino M 1987 CVD *in situ* ceramic composites. *J. Soc. Mater. Sci. Jpn.* **36**, 1205–11 (in Japanese)

Kawasaki A, Watanabe R 1999 Cyclic thermal fracture behavior and spallation life of PSZ/NiCrAlY functionally graded thermal barrier coatings. In: Kaysser W A (eds.) *Proc. 5th Int. Symp. Functionally Graded Materials '98*, Materials Science Forum Vols. 308–11. Trans Tech Publications, Switzerland, pp. 402–9

Li J-F, Watanabe R, Nishio N, Kawasaki A 1994 Design and fabrication of diamond tools with ceramic shank using the concept of functionally gradient materials. In: *Proc. Powder Metallurgy World Congress, PM '94,* Paris, pp. 553–6

Lin J S, Miyamoto Y, Tanihata K, Yamamoto M, Tanaka R 1998 Toughening effects of WC/Co particles and compressive surface stress on (Al_2O_3–WC/Co)/TiC/Ni graded materials. *J. Mater. Sci.* **33**, 869–76

Maruno S, Iwata H, Ban S, Itoh H 1999 Bone bonding studies and implant application of HA-G-Ti functionally graded composites. In: Kaysser W A (eds.) *Proc. 5th Int. Symp. Functionally Graded Materials '98*, Materials Science Forum Vols. 308–11. Trans Tech Publications, Switzerland, pp. 344–9

Miyamoto Y, Kaysser W A, Rabin B H, Kawasaki A, Ford R G 1999 *Functionally Graded Materials: Design Processing and Applications*. Kluwer, Dordrecht, The Netherlands

Neubrand A, Rödel J 1997 Gradient materials: an overview of a novel concept. *Z. Metallk.* **88**, 358–71

Sohda Y, Kude Y, Uemura S, Saitoh T, Wakamatsu Y, Niino M 1993 Carbon/carbon composites graded with SiC/C functionally gradient compositions. In: Holt J B, Koizumi M, Hirai T, Munir Z A (eds.) *Proc. 2nd Int. Symp. Functionally Graded Materials '92*, American Ceramic Society, Westerville, OH, pp. 125–32

Suresh S, Mortensen A 1998 *Fundamentals of Functionally Graded Materials: Processing and Thermomechanical Behaviour of Graded Metals and Metal–Ceramic Composites*. IOM Communications, London

Uchida T, Furukawa M, Kitano I, Koizumi K, Matsamura H 1970 Optical characteristics of a light-focusing fiberguide and its applications. *IEEE J. Quantum Electron.* **6**(10), 606–612

Yang Y Y 1998 Stress analysis in a joint with a functionally graded material under a thermal loading by using the Mellin transform method. *Int. J. Solids Struct.* **35** (12), 1261–87

Yang Y Y, Munz D, Schaller W 1997 Effect of the stress jump at the interface of a joint on the failure behavior. *Int. J. Fract.* **87**, L113–8

A. Neubrand
Technische Universität Darmstadt, Germany

This Page Intentionally Left Blank

G

Glass Fibers

The term glass or a glassy material represents a rather large family of materials with the common characteristic that their structure is noncrystalline. Thus, rigorously speaking, one can produce a glassy material from a polymer, metal, or ceramic. An amorphous structure is fairly common in polymeric materials. It is less so in metals, although metallic glass, generally in the form of a ribbon, can be produced by rapid solidification, i.e., by not giving enough time for crystallization to occur. In this article the focus is on silica-based inorganic glasses because of their great commercial importance as reinforcement fibers for polymer matrix composites.

1. Fabrication

The characteristic ability of glasses to be drawn stems from their viscoelastic nature, or in other words the fact that their deformation under stress is a strong function of time and deformation rate. Glassy materials, above the glass transition temperature, T_g, show Newtonian viscosity, i.e., the stress is linearly proportional to the strain rate. This property is exploited in the drawing of glass fibers. We can write, in terms of normal stresses and strains:

$$\sigma = \eta\dot{\varepsilon} \tag{1}$$

where σ is the stress, η is the viscosity, and $\dot{\varepsilon}$ is the strain rate. The deformation proceeds at a constant volume, i.e., an increase in length is accompanied by a decrease in the cross-sectional area. If an incremental increase in length is dl, and the corresponding decrease in cross-section is $-\mathrm{d}A$, then:

$$\frac{\mathrm{d}l}{l} = \mathrm{d}\varepsilon = -\frac{\mathrm{d}A}{A}$$

where l and A are the instantaneous values of length and cross-sectional area, respectively. We introduce the strain rate by dividing throughout by incremental time, dt. Thus:

$$\dot{\varepsilon} = \frac{\mathrm{d}\varepsilon}{\mathrm{d}t} = -\frac{\mathrm{d}A}{A}\frac{1}{\mathrm{d}t} = -\frac{\mathrm{d}A}{\mathrm{d}t}\frac{1}{A} = -\frac{\dot{A}}{A}$$

$$A\dot{\varepsilon} = -\dot{A} \tag{2}$$

If F is the applied load, then from Eqns. (1) and (2) we can write:

$$F = \eta\dot{\varepsilon}A = -\eta\dot{A} \tag{3}$$

Equation (3) states that the rate at which the fiber becomes thinner is proportional to the applied force *not the applied stress*. This means that thinner and thicker regions suffer cross-sectional reduction at an equal rate; hence, there is no tendency for preferential thinning of the drawn glass where the stress is locally raised by a local decrease in fiber cross-sectional area. Thinning of drawn viscous fibers is, thus, uniform and stable. This is different from tensile thinning of most materials, which becomes interrupted by rupture after a certain strain due to a preferential local thinning instability, known as necking.

The stability of glass fibers against localization of deformation, and hence against necking, is further reinforced by the fact that for a given load, as the viscosity increases the strain rate decreases. This has indeed very important implications in fiber drawing. In order to understand these implications we need to examine the temperature dependence of viscosity. This temperature dependence is given by:

$$\eta = A\exp(Q/RT)$$

where Q is the molar activation energy, A is a pre-exponential constant, and R is the universal gas constant. This equation indicates that the viscosity varies inversely in an exponential manner with the temperature. As thinner regions of the fiber cool more rapidly, the viscosity of such regions increases resulting in a decrease in the rate at which the thin region gets thinner. This allows the thicker regions to catch up and increases further the resistance of the drawn fiber material against necking.

2. Glass Fibers for Reinforcement of Polymers

A schematic representation of the conventional fabrication procedure for glass fibers is shown in Fig. 1. The raw materials are heated in a hopper and the molten glass is fed into electrically heated platinum or platinum–rhodium bushings having multiple holes. Typically, each bushing contains 200 holes at its base. A constant head of molten glass is kept in the tank. The molten glass flows by gravity through these holes forming fine continuous filaments which are gathered together and passed around a fast-rotating collet, followed by drawing at a speed of 1–2 km min^{-1}. Glass flows out of a platinum bushing at a rate given by:

$$q = \frac{khr^4}{\upsilon l}$$

where r is the radius, l is the length, υ is the kinematic viscosity (i.e., dynamic viscosity, η, divided by density, ρ), and h is the hydrostatic pressure generated by

---

* Cross references marked by an asterisk are included in this volume.

# Cross references marked by a hash can be found by consulting the Encyclopedia of Materials: Science and Technology.

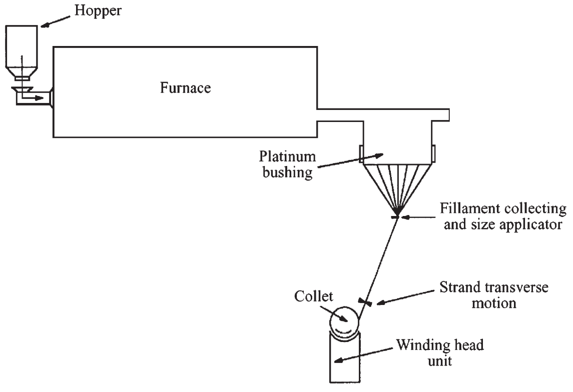


Figure 1
Schematic of the conventional fabrication procedure for glass fibers.

the head of glass in the bushing. The fibers exiting the bushing are cooled, subjected to a surface treatment (using an organic size such as a starch oil), and given a stretching treatment by the take-up spool. A size, consisting of an aqueous polymer emulsion, is applied before winding on a drum.

Typically, the glass fibers have a diameter between 5 μm and 20 μm. The final glass fiber diameter is a function of the bushing orifice diameter, viscosity (itself a function of composition and temperature), and the head of glass in the hopper. The viscosity is generally around 100 Pas for good fiberization. During the glass fiber stretching treatment by the take-up spool, the take-up speed is 50–60 ms^{-1}, which is slightly higher than the drawing speed, and thus the as-wound glass fiber suffers a stretch.

Glass filaments are easily damaged by the introduction of surface defects. The sizing treatment provides for ease of handling, protects against surface damage, and binds the filaments into a strand. In almost any fiber drawing process from a melt, the melt viscosity is by far the most important process control parameter. If a single glass filament breaks, the glass will drip across other filaments and cause a breakdown of the whole drawing process. Filament breakage can occur because of undissolved particles in the glass, abrasion, devitrification, or any instability due to changes in viscosity/winding speed (Proctor 1971).

An important characteristic of glass fiber is its extreme flexibility, which of course stems from its fine diameter and low modulus. Textile processes of weaving, knitting, etc. can be used with fine, flexible glass fibers to produce fabrics, tapes, sleeves, and ropes. E-glass fiber has a diameter between 3 μm and 20 μm and is commonly used for insulation purposes as well as for reinforcement of polymers. Frequently, glass fibers are chopped to obtain short, staple fibers.

3. *Sol–Gel Processing of Glass Fibers*

The conventional processing of glass fibers, as described above, involves direct melting of raw materials, followed by drawing of the molten glass through electrically heated platinum bushings. A major problem with any direct melting process is that the temperatures involved are very high. Quite frequently, the melt viscosity at reasonable temperatures is too high to allow proper homogenization. At times the oxides show large differences in volatility and direct melting results in large losses of some of the constituents. All these problems have led to processing of ceramics and glasses via a chemical route. This allows a high degree of homogeneity on a molecular scale and consequently high-purity glasses or ceramics can be obtained. In particular, the sol–gel technique of making glass fibers has become important.

Essentially the sol–gel route for making any glass or ceramic involves the formation of the appropriate glass or ceramic structure by chemical polymerization of suitable compounds in the liquid state (sol) at low temperatures. A sol is a colloidal suspension in which the individual particles are so small that no sedimentation occurs. The particle size in a sol generally varies between 1 nm and 100 nm. A gel is a suspension in which the liquid medium is viscous enough to behave more or less like a solid.

The precursor material for glass fibers in the form of a gel is obtained from the sol. The unwanted residues (water, organic compounds, etc.) are removed from the gel by heating at temperatures much lower than those used in direct-melting processes, and the desired glass or ceramic is obtained in an appropriate form (powder, film, fiber, etc.). As the solvent liquid is lost, the sol viscosity increases until it becomes rigid, i.e., it gels. A well-controlled rate of drying is required because of the large shrinkage that accompanies the drying process.

Fiber drawing is carried out in the course of sol-to-gel conversion. The gelled fibers are then heated to produce glass fibers. There are some very general but important points to be considered in sol and gel spinning. Spinning of dilute solutions into gels results in a minimum of polymer chain entanglement. Gelled fibers are then highly porous and elastic, and should be stretched by drawing while in a temperature gradient. This removes the solvent and decreases the porosity. The viscosity of sols and gels can also be a function of time at constant temperature; e.g., when a tetraethoxysilane [$\mathrm{Si(OC_2H_5)_4}$]–water–hydrochloric acid–alcohol solution of appropriate composition is held at near room temperature, hydrolysis and polycondensation occur and the solution viscosity increases. Fibers can be drawn at a viscosity of about 10Pas, which occurs during the course of sol-to-gel transition. It should be pointed out that for an alkoxide solution to be spinnable (Sakka 1982) the solution must have linear polymers and an appropriate ratio of $[H_2O]/[Si(OC_2H_5)_4]$, called the *r*-ratio. An *r*-ratio <2 gives a spinnable solution.

4. Chemical Composition

Silica-based glass fiber has been around for a long time. Common glass fiber is readily available commercially in a variety of different chemical compositions. Most glass fibers are silica based (~50–60% SiO_2) and contain a host of other oxides of calcium, boron, sodium, aluminum, iron, etc. They are commonly used for reinforcement of thermosetting and thermoplastic polymers. Table 1 gives the compositions of some commonly used glass fibers in composites. The designation E stands for electrical as E-glass is a good electrical insulator besides having good strength and a reasonable Young's modulus; C stands for corrosion as C-glass has a better resistance to chemical corrosion; S stands for high silica content and S-glass is able to withstand higher temperatures than others. Most continuous glass fiber produced is of the E-glass type.

5. Structure

Silica-based glass is a generic term representing an interesting and versatile class of materials. Glasses of various compositions can be obtained and they show very different properties. Structurally, however, all silica-based glasses have the same basic building block: a tetrahedron made up of four oxygen ions bonded covalently to a silicon ion at the center. The composition of a single tetrahedron is thus SiO_4. Each oxygen ion is, however, shared by *two* tetrahedrons, giving the bulk composition of SiO_2. It is this basic building block that is repeated in three dimensions in silica and silica-based materials. Different repeat patterns result in different structures for the same composition. The addition of other oxides of metals such as sodium, calcium, aluminum, magnesium, etc., to silica serves to alter the network structure and the bonding, and consequently the properties, mainly a reduction in viscosity which makes for easy processing. The presence of crystalline regions is undesirable in a glass because such regions act as stress raisers. Glass fibers are quite isotropic; Young's modulus and thermal expansion coefficient are the same along the fiber axis and perpendicular to it. This, of course, is a result of the amorphous, three-dimensional network structure of glass, i.e., there is no preferential alignment of any structural units such as polymer chains along the fiber axis.

6. Properties

Compared with most other inorganic fibers, silica-based glasses generally have a comparatively low density. The strength can be quite high, but the elastic modulus is not very high. Thus, while the strength-to-weight ratio of glass fibers is quite attractive in comparison with other inorganic fibers, their modulus-to-weight ratio is only moderate. Typical properties of different glass fibers are summarized in Table 2.

Although resistance to fire and many chemicals are very attractive features of glass, moisture can decrease the strength of glass fiber. In particular, the adsorption of moisture by freshly made glass fiber results in a drastic decrease in its strength. Glass fibers are also susceptible to what is called static fatigue, i.e., they cannot withstand loads for long periods of time. This is due to the phenomenon of stress corrosion induced by the ambient moisture.

Common glass fiber such as E-glass becomes severely corroded in an alkaline atmosphere such as in cement or concrete. Some special alkali-resistant glass fibers have therefore been developed. The properties of one such fiber, called Cemfil, are given in Table 2.

7. Fracture

Many brittle, amorphous materials such as thermosetting polymers and silica-based inorganic glasses show some tell-tale markings on their fracture surfaces (Mecholsky *et al.* 1977, Chandan *et al.* 1994). Typically, the fracture surface of a glass fiber shows four distinct regions. These regions surrounding the origin point of fracture are: a smooth mirror region, a mist region of small radial ridges, a hackle region

Table 1
Chemical composition of some common glass fibers.

Compound	E-glass	C-glass	S-glass	Cemfil
SiO_2	55.2	65.0	65.0	71.0
Al_2O_3	8.0	4.0	25.0	1.0
CaO	18.7	14.0		
MgO	4.6	3.0	10.0	
Na_2O	0.3	8.5	0.3	$Na_2O+K_2O=11$
K_2O	0.2			
Li_2O	7.3	5.0		<1.0

Table 2
Typical properties of some glass fibers.

Glass fiber type	Density ($g\,cm^{-3}$)	Tensile strength (GPa)	Young's modulus (GPa)
E-glass	2.54	1.7–3.5	69–72
S-glass	2.48	2.0–4.5	85
C-glass	2.48	1.7–2.8	70
Cemfil	2.70		80

consisting of large ridges, and finally a region of extensive crack branching. It turns out that the product of strength, σ, and the square root of the distance of each of these regions from the origin of fracture is a constant. One can write:

$$\sigma R_i^{0.5} = A_i$$

where R_i are the radii of the mirror-mist, mist-hackle, or crack-branching boundaries. These radii can be related to the initial flaw depth, a, or the half-width, b, through the fracture mechanics analysis:

$$\frac{c}{R_i} = \frac{K_{Ic} Y^2}{2A_i^2}$$

where $c = (ab)^{0.5}$, Y is a geometrical constant, and K_{Ic} is the critical stress intensity factor or fracture toughness.

8. Applications

Glass fiber is available in a variety of shapes and forms. Some of the important ones include continuous fiber and roving, staple fiber, and chopped strand mat. Staple fibers are strands of individual filaments 200–400 mm long and are excellent for providing bulkiness for filling, filtration, etc. Chopped strand glass fibers consist of fibers chopped to various lengths, 3–50 mm, mainly for the purpose of mixing with a resin for making composites. Glass fiber mats consist of randomly dispersed chopped fibers or continuous fiber strands, held together with a resin.

Glass fibers find applications in a vast array of markets.

8.1 Automotive Market

The automobile industry is one of the largest users of glass fiber. Polymer matrix composites containing glass fibers are used to make external body panels, bumper beams, pultruded body panels and air ducts, engine components, etc. Parts made of glass-fiber-reinforced polymers are much lighter than metallic ones, making the automobile more fuel efficient.

8.2 Aerospace Market

Glass-fiber-reinforced composites are used to make aircraft parts such as wings, helicopter rotor blades, engine ducts, etc. Glass fiber has a relatively low elastic modulus. Hence, it is more common to use glass-fiber-reinforced polymer composites in the interior of an airplane rather than in primary structural parts. The radar transparency characteristic of glass has given it some key uses in the radar-evading stealth technologies.

8.3 Marine Market

Sailing boats and hulls and decks of commercial fishing boats and military mine-hunters are frequently made of glass-fiber-reinforced polymers. Glass-fiber-reinforced polyester is commonly used in making boats of all sizes.

8.4 Civil Construction

This is another large market where glass fiber is widely used. Typical applications include the use of glass fibers in polymeric resins for paneling, bathtubs and shower stalls, doors, windows, etc. Glass fibers are also used as reinforcement in a variety of household items such as paper, tapes, lampshades, etc. Some special alkali-resistant glass fibers have been developed for reinforcement of cement and concrete. Commonly, steel bars are used for such purposes. Cement, however, is very alkaline. An ordinary glass fiber such as E-glass will be severely corroded in an alkaline atmosphere, hence the need for special, alkali-resistant glass fibers (Majumdar 1970, Hannant 1978). The bonding between glass fiber and cement is mainly mechanical and has its origin in the shrinkage that occurs when the cement is set.

8.5 Sporting Goods

Not surprisingly, the sporting goods industry was one of the first to make use of glass-fiber-reinforced composites. Examples include bicycle frames, tennis rackets, golf-club shafts, cricket bats, skis, etc. Braided fibers in a resin matrix give high torsional stiffness to skis.

8.6 Electrical/Electronic Market

Glass fibers are used extensively in printed circuit boards, industrial circuit breakers, conduits for power cables, etc.

9. Conclusions

Glass fibers are a very versatile class of materials. They are used extensively as a reinforcement fiber for polymeric resins such as epoxy and unsaturated polyester. The stiffness of glass fiber is lower than that of other reinforcement fibers, but it possesses the distinct advantage of combining a very high strength with low density and, most of all, a very reasonable cost. Glass fiber will continue to be used as a major reinforcement fiber well into the future.

Bibliography

Chandan H C, Parker R D, Kalish D, Bradt R C 1994 In: Tressler R E (eds.) *Fractography of Glass*. Plenum, New York, p. 143

Hannant D J 1978 *Fibre Cements and Fibre Concretes*. Wiley, New York

Majumdar A J 1970 Glass fibre-reinforced cement and gypsum products. *Proc. R. Soc. London* **A319**, 69

Mecholsky J J, Freiman S W, Morey S N 1977 Fractographic analysis of optical fibres. *Bull. Am. Ceram. Soc.* **56**, 1016

Proctor B A 1971 Fibre-reinforced composite materials. *Composites* **2**, 85

Sakka S 1982 In: *Treatise on Materials Science & Technology, 22*. Academic Press, New York, p. 129

K. K. Chawla

University of Alabama, Birmingham, Alabama, USA

This Page Intentionally Left Blank

H

Hard Tissues, Mechanical Properties of

Bone has a varied arrangement of material structures at many length scales, which work in concert to perform diverse mechanical, biological, and chemical functions such as, structural support, protection and storage of bone cells, and mineral ion homeostasis. In order to understand the mechanical properties of hard tissues, it is important to understand the mechanical properties of its component phases, and the structural relationship between them at the various levels of hierarchical structural organization (Ziv *et al.* 1996). For example, a piece of writing chalk and a clamshell are chemically almost identical in composition, but the chalk will break far more easily. The difference in the performance of these two materials originates in their microstructure. This hierarchically organized structure has an irregular, yet optimized, arrangement and orientation of the components, making the material of bone heterogeneous and anisotropic (Fig. 1).

It has been shown that the mechanical properties of bone vary at different structural levels. For example, the Young's modulus of large tensile cortical specimens has been shown to be in the 14–20 GPa range (Reilly *et al.* 1974) (wet specimen, macrostructural property), while that of micro-bending cortical specimens was 5.4 GPa (wet specimen, microstructural property; Choi *et al.* 1990). However, it is unclear whether this discrepancy is due to the testing method or due to the influence of microstructure. Recently, the Young's modulus of osteon lamellar bone measured by nanoindentation was approximately 22 GPa (dry specimen, submicrostructure property; Rho *et al.* 1997) close to that for the macrostructure. There are, at present, no data in terms of Young's modulus for the nanostructure. The basic building blocks (apatite crystals and collagen fibrils) are extremely small, making mechanical testing nearly impossible. Therefore, it seems of primary importance to break down the mechanical testing of bone according to the various levels or architecture within bone material.

1. Macrostructural Properties: Cortical and Cancellous Bone

Some workers (Carter and Hayes 1977, Keller *et al.* 1990) consider cortical and cancellous bone to be a single morphological material, which can be characterized by a highly variable porosity or apparent density (mass/total volume, including pores). Others consider these two types to consist of two different materials (Choi *et al.* 1990, Rho *et al.* 1993). In general, cancellous bone material is much more active metabolically, is remodeled more often than cortical bone, and is, therefore, "younger" on average than cortical bone. Therefore, even though cancellous and cortical bone may be of the same kind of material, the maturation of the cortical bone material may alter the mechanical properties at the microstructural level.

At the macrostructure level, bone is distinguished into the cortical (or compact) and cancellous (or

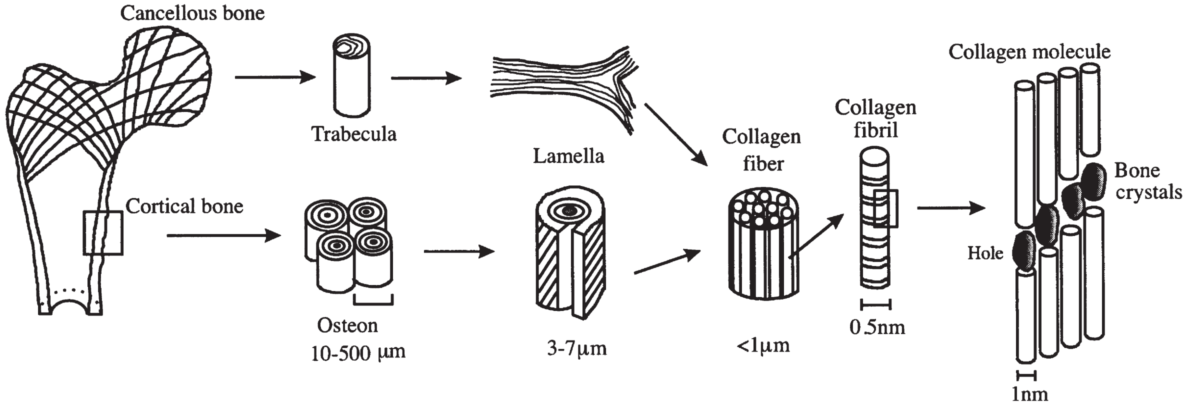


Figure 1
Hierarchical structural organization of bone: (a) macrostructure—cortical and cancellous bones, (b) microstructure—osteons and trabeculas, (c) sub-microstructure—lamellas, (d) nanostructure—mineralized collagen fibrils, (e) sub-nanostructure—mineral, collagen, and noncollagenous organic proteins.

---

* Cross references marked by an asterisk are included in this volume.

# Cross references marked by a hash can be found by consulting the Encyclopedia of Materials: Science and Technology.

trabecular) types. In cross-section, the end of a long bone such as the femur has a dense cortical shell with a porous, cancellous interior. Flat bones such as the calvaria have a sandwich structure composed of dense cortical layers on the outer surfaces and a thin, reinforcing cancellous structure within. Although both types of bone (cortical and cancellous) are most easily distinguished by their degree of porosity or density, true differentiation comes from histological evaluation of tissue microstructure.

The mechanical properties of bone are influenced greatly by the porosity and the organization of the microstructure. It is, therefore, difficult to predict micro-properties *in vivo* by measuring mechanical properties at the macrostructural level. In general, values for mechanical properties of bone at the macrostructural level vary from one bone to another as well as within different regions of the same bone (Rho *et al.* 1995, Goldstein 1987). The lesser regional heterogeneity observed in cortical bone is the result of the lower turnover rate for this material, which engenders changes at the microstructural level. Overall, the differences between the mechanical properties in cancellous bone are much broader than those in cortical bone and indeed they may vary by a factor of 2–5 from bone to bone. To a great extent, differences between whole bones may be explained by the variance in their structure and mechanical function (Burstein *et al.* 1976, Yamada 1970, Cezayirlioglu *et al.* 1985, Martens *et al.* 1983, Linde *et al.* 1989, Keaveny *et al.* 1997). However, the properties of a particular type of bone can not be adequately expressed in single values; our best estimates can only be given in a range of values reflecting both the experimental difficulties and the observed regional heterogeneity (Table 1).

2. *Microstructural Properties: Osteons and Trabeculas*

Mineralized collagen fibers form into planar arrangements called *lamellas* (3–7 μm wide). In some cases these sheets (lamellas) of mineralized collagen fibers wrap in concentric layers (3–8 lamellas) around a central canal to form what is known as an *osteon*. The osteon looks like a cylinder about 200–250 μm in diameter running roughly parallel to the long axis of the bone. Other forms of cortical bone, where the mineralized collagen fibers are less well registered and no pattern can be distinguished, are called *woven* bone. In some forms of bone, the lamellas are overall tangential to the outer surface of the bone (without forming osteons), and together with woven bone tissue, form a larger plywood-type stacking of thicker layers (150–300 μm) around the complete circumference of the bone. Cancellous bone is made of an interconnecting framework of trabeculas in a number of combinations all comprising of the following basic cellular structures: rod–rod, rod–plate, or plate–plate. A trabecular rod is about 50–300 μm in diameter.

Over the years Ascenzi and co-workers have examined the mechanical properties of single Haversian systems in bending (Ascenzi *et al.* 1990), tension (Ascenzi and Bonucci 1967), compression (Ascenzi and Bonucci 1968), and torsion (Ascenzi *et al.* 1994; Table 2). Osteons with longitudinal lamellas are better for tension and torsion and perhaps stronger in bending as well, while osteons with alternating lamellas are more suited for compression. The stiffnesses produced from torsion testing of individual osteons were much higher than those for the whole bone. Ascenzi *et al.* do not justify or explain this

Table 1
Mechanical properties of human cortical and cancellous bones tested by compression, tensile, and torsional testing.

Bone	Specimen		Strength (MPa)	Elastic modulus
Human cortical bone	Femur[a]	Compression	179–209	15.4–18.6 GPa
	Tibia[a]		183–213	24.5–34.3 GPa
	Femur[a]	Tensile	120–140	15.6–17.7 GPa
	Tibia[a]		45–170	18.9–29.2 GPa
	Femur[b]	Torsional	54 ± 0.6	3.2 GPa
	Tibia[c]		66–71	–
	Femur[d]	Bending	103–238	9.82–15.7 GPa
Human cancellous bone	Femoral head[e]	Compression	9.3 ± 4.5	900 ± 710 MPa
	Femoral neck[e]		6.6 ± 6.3	616 ± 707 MPa
	Intertrochanteric[e]		3.6 ± 2.3	263 ± 170 MPa
	Proximal tibia[f]		5.3 ± 2.9	445 ± 257 MPa
	Vertebral body[g]		–	165 ± 110 MPa

[a] Burstein *et al.* 1976, [b] Yamada 1970, [c] Cezayirlioglu *et al.* 1985, [d] Keller *et al.* 1990, [e] Martens *et al.* 1983, [f] Linde *et al.* 1989, [g] Keaveny *et al.* 1997.

result. However, this effect may be explained by the use of Cosserat micropolar elasticity, which allows for local rotation at cement sheaths and other internal boundaries (Lakes 1995). Katz (1981) confirmed that there is a majority of fibers in the longitudinal direction in osteons and also along the long axis of trabeculas in cancellous bone, but the situation in general is muddled. Marotti (1993) claims that fibers in general follow two patterns which constitute thin and thick lamellas; the thin ones are more oriented and compact, the thick ones are more diverse and sparse (somewhat microporous) in their elements.

The trabecular properties are easier to study in isolation. However, in spite of recent reports (Choi *et al.* 1990, Rho *et al.* 1993) there remains some controversy regarding the value of the elastic modulus of single trabeculas (Table 3). Trabecular bone material properties are important for characterizing various bone pathologies, and the remodeled bone adjacent to various joint implants, because they are affected by disease sooner than cortical bone. In the past it was assumed that individual trabeculas, single osteons, and a thin cortical shell possessed the same mechanical properties as those of large cortical bone specimens regardless of their type or size (Wolff 1892). However, many investigators produced values for the elastic modulus of individual trabeculas, single osteons, and a thin cortical shell that were considerably less than that for whole bone (Choi *et al.* 1990, Rho *et al.* 1993, Ashman and Rho 1988). Mechanical testing of cortical osteonal samples has shown a similar modulus, while individual trabeculas show a lower modulus of elasticity than that of large cortical bone specimens. This issue needs further clarification.

A literature survey of measured and estimated values of the modulus of trabecular bone material shows that moduli values range from 1 GPa to 20 GPa (Table 3). Rho *et al.* (1993) showed that the relationship derived from this data, between elastic moduli and density in cancellous bone material, could not be extrapolated from similar data from

Table 2
Mechanical properties of single osteons and trabeculas. Longitudinal (L) osteon has marked longitudinal spiral course of fiber bundles in successive lamellas. Alternate (A) osteon has fiber bundles in one lamella making an angle of nearly 90° with the bier bundles of the next one.

Test method	Elastic modulus (GPa)
L osteon, three-point bending[a]	2.32 ± 1.20
A osteon, three-point bending[a]	2.69 ± 0.93
L osteon, tension[b]	11.7 ± 5.8
A osteon, tension[b]	5.5 ± 2.6
L osteon, compressive[c]	6.3 ± 1.8
A osteon, compressive[c]	7.4 ± 1.6
T osteon, compressive[c]	9.3 ± 1.6
L osteon, torsion[d]	3.3 ± 0.5
A osteon, torsion[d]	4.1 ± 0.4
T osteon, torsion[d]	4.1 ± 0.4

[a] Ascenzi *et al.* 1990, [b] Ascenzi and Bonucci 1967, [c] Ascenzi and Bonucci 1968, [d] Ascenzi *et al.* 1982.

Table 3
Mechanical properties of single trabeculas.

Test method	Elastic modulus (GPa)
Hypothesis[a]	17–20 (assumption)
Buckling[b]	8.69 ± 3.17 (dry)
Inelastic buckling[c]	11.38 (wet)
Back-calculating from finite element models[d]	1.30
Ultrasound test method[e]	12.7 ± 2.0 (wet)
Tensile testing[f]	0.76 ± 0.39
Microhardness[g]	15 (estimation)
Three-point bending[h]	3.81 (wet)
Cantilever bending with finite element analysis[i]	7.8 ± 5.4 (dry)
Four-point bending[j]	5.35 ± 1.36 (wet)
Tensile testing[k]	10.4 ± 3.5 (dry)
Ultrasound test method[k]	14.8 ± 1.4 (wet)
Nanoindentation[l]	19.6 ± 3.5 (dry): longitudinal direction 15.0 ± 3.0 (dry): transverse direction

[a] Wolff 1892, [b] Runkle and Pugh 1975, [c] Townsend *et al.* 1975, [d] Williams and Lewis 1982, [e] Ashman and Rho 1988, [f] Ryan and Williams 1989, [g] Hodgskinson *et al.* 1989, [h] Kuhn *et al.* 1989, [i] Mente and Lewis 1989, [j] Choi *et al.* 1990, [k] Rho *et al.* 1993, [l] Rho *et al.* 1997.

tests on cortical bone material and its density. This led the authors to conclude that the materials of the two bones were intrinsically different. However, the data are not definitive, as the dependence of the modulus on other parameters such as location, microstructure, or density variation was not considered. The most recent studies by use of nanoindentation (Rho *et al.* 1999) and by finite element analysis (FEA) simulation (vanRietbergen *et al.* 1996) suggest that in fact the elastic properties of single trabeculas are very similar to the properties of nearby cortical tissue. This matter requires further clarification because imprecise values may lead to misinterpretations of the structural function of each type of bone material or misinterpretation of the role of trabecular bone in the mechanical behavior of normal and implanted joints.

3. Submicrostructural Properties: Lamellas

Bone lamellas are 3–7 μm thick, but the detailed anatomy of lamellas is still a matter of dispute. Gebhardt (1906) all lamellas are dense fibrous layers in which collagen fibers lie parallel; the difference between successive lamellas consists only in the orientation of the fibrils, which may vary through an angle of 0° to 90°. Ziv *et al.* (1996) proposed a model in which adjacent lamellas have roughly the same structure, but their layers are aligned in different orientations. There are transitional regions between the neighboring lamellas in which the orientation of the collagen and its associated mineral changes rapidly but smoothly over a few microns. In other models, the lamellar pattern of the bone is held to depend on the alternation of layers of heterogeneous structure (Reid 1986). The models propose that lamellar bone is made up of alternating collagen-rich (thin lamellas) and collagen-poor (thick lamellas) layers, all having an interwoven arrangement of fibers. By means of x ray microprobe analysis, Marotti (1993) found that the calcium and phosphorus content of thick lamellas was 10–15% higher than in thin lamellas. Reid (1986) also observed that thin lamellas are more rapidly removed during polishing than thick lamellas.

With nanoindentation, it is possible to measure properties at the micron and submicron levels and thus study the properties of bone microstructural components having dimensions of only a few microns or less. For example, nanoindentation has recently been used to determine the elastic modulus and hardness of individual trabecular struts (Rho *et al.* 1997, 1999) and small filler particles in resin composites and other dental restoratives (Townsend *et al.* 1975). The nanoindentation technique is theoretically able to measure mechanical properties with a resolution of better than 1 μm. To the best of our knowledge, in only one previous study has an investigation on the tensile properties of individual lamellas within osteons been carried out (Ascenzi *et al.* 1982). It met with little success due to the inhomogeneity in thickness of single lamellar samples along their length and width. They observed that there was usually a wide, continuous range of variation in thickness even along the length of single osteonic lamellas. However, the dimensions of these substructural building blocks of osteons, a few microns or less, were well within the capabilities of nanoindentation measurement.

The mechanical properties of the osteonal thick lamellas (19.7 GPa) have been successfully measured using nanoindentation. However, the properties of thin lamellas, which some investigators (Gebhardt 1906, Ziv *et al.* 1996) believe to have a quite different microstructure compared with that of the thick lamellas, still remain to be measured in order to gain a better understanding of the properties of lamellar bone.

4. Nanostructural Properties: Collagen Fibers and Crystals

The most prominent structures seen at this scale are the collagen fibers, surrounded and infiltrated by mineral. The attachment sites of macromolecules onto the collagen framework are not distinctly known, although several immunohistological studies have shown preferential labeling of some macromolecules in a periodic fashion along the collagen molecules and fibers (Glimcher 1989). The subnanostructures of the three main materials are crystals, collagens, and noncollagenous organic proteins.

Some studies, which used selective demineralization and acoustic methods, have come up with most intriguing theories about the complementary role of the collagen and the mineral (Turner *et al.* 1995, Pidaparti *et al.* 1996). These workers have suggested that isolated collagen is more or less isotropic and by the impregnation of mineral it reaches the anisotropic ratios that are known for whole bone (1.7–2.1) in two normal directions. Experimental methods that can resolve the issue of absolute and relative (anisotropy) values for the elastic modulus of microscopic bone tissue would be invaluable. There is still no reliable data regarding the elastic properties of the mineral crystals and the collagen matrix of bone. Isolated crystals retaining all of the native features of bone mineral can be obtained using low-temperature and nonaqueous methods; however, it is not clear how the mechanical properties of these crystals can be measured at this time. Most likely, the elastic properties of bone mineral crystals are different from those of pure hydroxyapatite due to the vacancies and substitutions in the natural crystals. Examination of the highly mineralized rostrum bone of a Mesoplodon whale (it has only 2–3% collagen w/w) shows that bone fully compacted with mineral retains the same anisotropy ratios (roughly 2:1 in elasticity and toughness) as less mineralized "normal" bone. This seems to underline the importance of the shape, geometry, and material properties of the crystals for determining the behavior of whole bone.

The elastic properties of the collagen of bone have been assumed to be the same as that for tendon or ligament collagen. However, unlike the collagen of soft tissues, it is normal to assume that bone collagen will not show a reorientation under stretch and also will not (does not) have the crimp pattern and straightening out effect. Examining the organic remnant after crystal dissolution appears to be the customary method for obtaining the properties of the collagen of bone (Bowman *et al.* 1996). However, no one can be certain of the constitution and any conditional changes of the phase (either mineral or collagen) that is left behind (Aerssens *et al.* 1994). Therefore, it is very difficult to imagine what the bone-collagen properties may be, and certainly the intimate relationship of mineral and collagen and their interplay makes measurement and predictions even more difficult. Further investigations are warranted for a more complete understanding as to how mineral–matrix interactions are related to their effects on mechanical properties at the various hierarchical levels of tissue structure and organization.

5. *Concluding Remarks*

Additional research would advance our understanding of bone as a structure and as a material. The following questions remain unanswered: Is there a size effect in bone? What is the hierarchical level at which homogenized material properties can be obtained that are similar to the ones observed macroscopically in whole bone samples? What is the nature of the interfaces between the constituents of bone at all length scales? If we are successful in answering these questions, various pathophysiological conditions may be easier to understand. Aging, osteoporosis, osteoarthritis, and other degenerative diseases all are expected to influence, and be influenced by, the microstructure to different degrees. It may just be possible that certain diseases affect the constitution and quality of only one level of the hierarchical organization of bone, and, therefore, they may become easier to understand and treat. Therefore, further investigations of mechanical properties at the "materials level" in addition to the studies at the "structural level" are needed to fill the gap in our present knowledge and to achieve a better understanding of the mechanical properties of hard tissues such as bone.

See also: Bone and Natural Composites: Properties*; Collagen Derived Materials#

Bibliography

Aerssens J, Dequeker J, Mbuyi-Muamba J-M 1994 Bone tissue composition: Biochemical anatomy of bone. *Clin. Rheumat.* **13**, 54–62 (Suppl. 1)

Ascenzi A, Baschieri P, Benvenuti A 1990 The bending properties of single osteons. *J. Biomech.* **23**, 763–71

Ascenzi A, Baschieri P, Benvenuti A 1994 The torsional properties of single selected osteons. *J. Biomech.* **27**, 875–84

Ascenzi A, Benvenuti A, Bonucci E 1982 The tensile properties of single osteonic lamellae: technical problems and preliminary results. *J. Biomech.* **15**, 29–37

Ascenzi A, Bonucci E 1967 The tensile properties of single osteons. *Anat. Rec.* **158**, 375–86

Ascenzi A, Bonucci E 1968 The compressive properties of single osteons. *Anat. Rec.* **161**, 377–92

Ashman R B, Rho J Y 1988 Elastic moduli of trabecular bone material. *J. Biomechanics* **21**, 177–81

Bowman S M, Zeind J, Gibson L J, Hayes W C 1996 The tensile behavior of demineralized bovine cortical bone. *J. Biomech.* **29**, 1497–501

Burstein A H, Reilly D T, Martens M 1976 Aging of bone tissue: mechanical properties. *J. Bone Jt Surg.* **58A**, 82–6

Carter D R, Hayes W 1977 The compressive behavior of bone as a two-phase porous structure. *J. Bone Jt Surg.* **59A**, 954–62

Cezayirlioglu H, Bahniuk E, Davy D T, Heiple K G 1985 Anisotropic yield behavior of bone under combined axial force and torque. *J. Biomech.* **18**, 61–9

Choi K, Kuhn J L, Ciarelli M J, Goldstein S A 1990 The elastic moduli of human subchondral trabecular, and cortical bone tissue and the size-dependency of cortical bone modulus. *J. Biomech.* **23**, 1103–13

Gebhardt W 1906 Über funktionell wichtige Anordnungsweisen der feineren und gröberen Bauelemente des Wirbeltierknochens, II. Spezieller Teil. Der Bau der Haversschen Lamellensysteme und seine funktionelle Bedeutung. *Arch. Entwickl. Mech. Org.* **20**, 187–322

Glimcher M 1989 Mechanisms of calcification: role of collagen fibrils and collagen-phosphoprotein complexes *in vitro* and *in vivo*. *Anat. Rec.* **224**, 139–53

Goldstein S 1987 The mechanical properties of trabecular bone: dependence on anatomic location and function. *J. Biomech.* **20**, 1055–61

Hodgskinson R, Currey J D, Evans G P 1989 Hardness, an indicator of the mechanical competence of cancellous bone. *J. Orthop. Res.* **7**, 754–8

Katz L J 1981 Mechanical properties of bone. *ASME AMD* **45**, 171–84

Keaveny T M, Pinilla T P, Crawford R P 1997 Systemic and random errors in compression testing of trabecular bone. *J. Orthop. Res.* **15**, 101–10

Keller T S, Mao Z, Spengler D M 1990 Young's modulus, bending strength, and tissue physical properties of human compact bone. *J. Orthop. Res.* **8**, 592–603

Kuhn J L, Goldstein S A, Choi K W, London M, Feldkamp L A, Matthews L S 1989 Comparison of the trabecular and cortical tissue moduli form human iliac crests. *J. Orthop. Res.* **7**, 876–84

Lakes R 1995 On the torsional properties of single osteons. *J. Biomech.* **28**, 1409–10

Linde F, Hvid I, Pongsoipetch B 1989 Energy absorptive properties of human trabecular bone specimens during axial compression. *J. Orthop. Res.* **7**, 432–9

Marotti G 1993 A new theory of bone lamellation. *Calcif. Tissue Int.* **53**(Suppl. 1), S47–S56

Martens M, Van Audekercke R, Delport P 1983 The mechanical characteristics of cancellous bone at the upper femoral region. *J. Biomech.* **12**, 971–83

Mente P L, Lewis J L 1989 Experimental method for the measurement of the elastic modulus of trabecular bone tissue. *J. Orthop. Res.* **7**, 456–61

Pidaparti R M, Chandran A, Takano Y, Turner C H 1996 Bone mineral lies mainly outside collagen fibrils: predictions of a composite model for osteonal bone. *J. Biomech.* **29**, 909–16

Reid S A 1986 A study of lamellar organization in juvenile and adult human bone. *Anat. Embryol.* **174**, 329–38

Reilly D T, Burstein A H, Frankel V H 1974 The elastic modulus of bone. *J. Biomech.* **7**, 271–5

Rho J Y, Ashman R B, Turner C H 1993 Young's modulus of trabecular and cortical bone material: ultrasonic and microtensile measurements. *J. Biomech.* **26**, 111–9

Rho J Y, Hobatho M C, Ashman R B 1995 Relations of density and CT numbers to mechanical properties for human cortical and cancellous bone. *Med. Eng. Phys.* **17**, 347–55

Rho J Y, Roy M E, Tsui T Y, Pharr G M 1999 Elastic properties of microstructural components of human bone measured by nanoindentation. *J. Biomed. Mater. Res.* **45**, 48–54

Rho J Y, Kuhn-Spearing L, Zioupos P 1998 Mechanical properties and the hierarchical structure of bone. *Med. Eng. Phys.* **20**, 92–102

Rho J Y, Tsui T Y, Pharr G M 1997 Elastic properties of human cortical and trabecular lamellar bone measured by nanoindentation. *Biomaterials* **18**, 1325–30

Runkle J C, Pugh J 1975 The micro-mechanics of cancellous bone. *Bull. Hosp. Jt Dis.* **36**, 2–10

Ryan S D, Williams J L 1989 Tensile testing of rodlike trabeculae excised from bovine femoral bone. *J. Biomech.* **22**, 351–5

Townsend P R, Rose R M, Radin E L 1975 Buckling studies of single human trabeculae. *J. Biomech.* **8**, 199–201

Turner C H, Chandran A, Pidaparti R M V 1995 The anisotropy of osteonal bone and its ultrastructural implications. *Bone* **17**, 85–9

vanRietbergen B, Odgaard A, Kabel J, Huiskes R 1996 Direct mechanics assessment of elastic symmetries and properties of trabecular bone architecture. *J. Biomech.* **29**, 1653–7

Weiner S, Traub W 1992 Bone structure: from angstroms to microns. *FASEB* **6**, 879–85

Williams J L, Lewis J L 1982 Properties and an anisotropic model of cancellous bone from the proximal tibial epiphysis. *J. Biomech. Eng.* **104**, 50–6

Wolff J 1892 *Das Gesetz der Transformation der Knochen*. Hirschwald, Berlin

Yamada H 1970 *Strength of Biological Materials*. Williams & Wilkins, Baltimore, MD

Ziv V, Wagner H D, Weiner S 1996 Microstructure-microhardness relations in parallel-fibered and lamellar bone. *Bone* **18**, 417–28

J.-Y. Rho
University of Memphis, Tennessee, USA

High-performance Fibers

1. *Fiber Development*

High-performance synthetic fibers often resemble textile fibers in appearance and some are of the same chemical composition; however, they are produced so as to optimize their mechanical properties to achieve enhanced performance (Bunsell A R 1988, Hearle J W S 2001, and Moncrieff R W 1979). Such fibers are used in structures for which their flexibility, lightness, and very high mechanical properties, compared to bulk materials, make them indispensable. These fibers have higher strengths and often higher elastic moduli than conventional textile fibers due to the manufacturing processes employed and also, in many cases, due to their molecular or atomic structures. In the case of some high-performance fibers, their properties can attain almost the limits of what is physically possible so that very high moduli, even up to that of graphite, can be achieved. Some fibers have been developed to possess great thermal stability for use at temperatures above the limits at which high-performance metals can be used. Although stiff along their lengths, fibers are flexible when bent because of their small diameters, which are usually of the order of 10 μm. They are used in applications that include high-performance tethering ropes, flexible architectural structures, reinforced elastomers, and fiber-reinforced composites.

The commercial development of synthetic fibers began in the 1930s with the production of glass fibers in the US. Glass and glass fibers have approximately the same Young's modulus and density as that of aluminum. Although fibers made from regenerated cellulose, from wood, or other plants, were produced, first in France, toward the end of the nineteenth century and are known as rayon or viscous rayon, the first truly synthetic polymeric fibers were produced in 1938. These were polyamide 6.6 (developed in the US) and polyamide 6 fibers (developed in Germany). These polyamide, or "nylon" fibers, together with polyester fibers, developed in Great Britain in the 1940s, are the most important synthetic textile fibers produced. Although of low stiffness, they were the first of a family of organic fibers, which have been developed with elastic moduli rivaling and even exceeding that of steel, but with only a fifth or less of the weight. These organic fibers find, however, many industrial applications as in tyre reinforcement and anchor lines for off-shore oil platforms. Boron and silicon carbide fibers made by chemical vapor deposition onto a filament of tungsten or carbon were first produced in the early 1960s in the US and then in France and Russia (Kostikov VI 1995). They have the density of glass but Young's modulus twice that of steel. Carbon fibers, with a density of a quarter of that of steel, were first produced in Great Britain in the mid-1960s. They are made by the pyrolysis of polyacrylonitrile precursor fibers and commonly have moduli up to one and half times the modulus of steel but can be up to three times stiffer. An alternative process developed in the 1970s in the US uses the pitch residue from oil refining or the coking of coal to give fibers with moduli approaching that of graphite. Fine ceramic fibers, for structural applications, started to appear at the end of the 1970s and in the 1980s (Bunsell A R and Berger M-H 1999). One group of these fibers, developed in Great Britain and in the US, is based on alumina, whereas others developed in Japan are based on silicon carbide. These fibers have Young's moduli ranging from that of steel to twice this value with densities less than half that of steel.

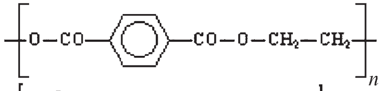
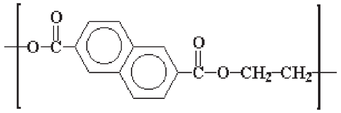
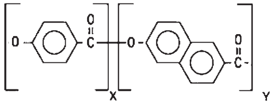
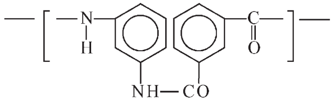
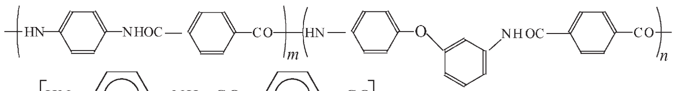
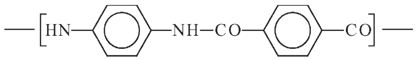
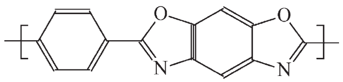
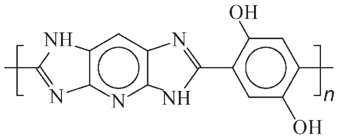
Most fibers have diameters of the order of 10 μm, which can be compared to the diameter of a human hair, which is 70 μm. Exceptions, however, are the boron and SiC fibers made by Chemical vapor deposition (CVD) onto cores, which are twice as thick as hair. The fineness of fibers allows even the stiffest of materials to be made in a flexible form as bending stiffness is related to the cube of the diameter. However, many high-modulus fibers are completely elastic, at least at room temperature, and are therefore brittle in tension and in bending. They are most suited to be used as reinforcements in composite materials. The high-performance polymer fibers are also used as reinforcements, such as in tyres and other composites, but they are not brittle as they deform plastically in compression. They show great toughness which can be exploited (Chawla K K 1998) in structures such as cables and structures, which must resist impacts.

1.1 Organic Fibers

Polyamide (PA) or "nylon" fibers were developed in 1938 by Du Pont in the US and in Germany by I.G. Farben. These fibers are based upon amide groups (–C(O)NH–) contained within a chain of carbon atoms making up a backbone of atoms linked by strong covalent bonds (Table 1). However, the molecules are linear and can bend and rotate at each bond so that the molecules are not perfectly aligned parallel to the fiber axis, which results in a low Young's modulus. Around 1947, polyethylene terephthalate (PET) or "polyester" fibers were produced in Great Britain and have become the most widely produced synthetic fibers. Table 1 shows that the polyester molecule contains an aromatic ring, which stiffens the molecule compared to the very flexible polyamide structure. Both polyamide and polyester fibers melt at around 260 °C. Although both of these latter fibers melt at around the same temperature, it is significant that the polyester fiber, which contains the aromatic rings, shows much higher resistance to discoloration at raised temperatures than the polyamide fiber. A fiber which is related to PET fibers is the polyethylene-naphthalate (PEN) fiber. Table 1 shows that this latter fiber is based on a molecule which has two aromatic rings, which further increases the rigidity of the fiber. The three fibers mentioned above are made by drawing from the melt through a spinneret. The drawing rate determines the final properties. An undrawn fiber will have a strain to failure of some hundreds of a percent and tensile loading will produce not only a very high strain to failure but deformation will involve striction of the filament. Striction is when the diameter of the filament is locally reduced under load and in this region the molecular structure becomes aligned parallel to the fiber axis and the loading direction. Further loading induces an increase in the striction until the specimen is reduced in cross section and the molecular structure is largely oriented parallel to the fiber axis. Fibers for textile end-uses are drawn during production so as to produce filaments with enhanced molecular orientation, higher strengths, and longitudinal stiffness but with reduced strains to failure, generally of the order of 30%. Striction is not usually seen with this type of fiber. High-performance versions of PA and PET fibers, which are utilized for industrial applications, are more highly drawn with further reduced strains to failure in the region of 20% for the former and 15% for the latter. PEN fibers are only produced as high-performance industrial fibers and have strains to failure of around 6%. The molecular morphology of these fibers consists of microfibrils, with transverse dimensions of around 5 nm, arranged more or less parallel to the fiber axis. Regions within the microfibrils are composed of well-ordered arrangements of molecules which are folded as in a concertina. The rest of the structure is made up of less well-ordered regions so that the overall molecular structure is semi-crystalline with a crystalline content generally of the order of 40%.

A typical tensile fracture morphology of one of the above thermoplastic fibers is shown in Fig. 1(a). It can be seen that the fracture surface consists of two regions, an initial inclined region corresponding to initiation at the surface and slow crack propagation normal to the fiber axis. The crack opens because of plastic deformation ahead of the crack. Rapid failure finally occurs when the remaining part of the section can no longer support the applied stress. Under particular cyclic conditions, the same fiber fails in fatigue and one half of the fatigue break is shown in Fig. 1(b). The crack has initiated at or near the surface but instead of running normal to the fiber axis it has been deviated to run at a slight angle to the fiber axis, penetrating into the fiber section so that the remaining section finally fails by the tensile process. A particularity of this type of failure is that the minimum load has to be lower than a threshold level so that

Table 1
Molecular composition of high-performance organic fibers.

Fiber type	Repeat unit in the macromolecule	Maximum elastic modulus (GPa)	Melting or decomposition temperature (°C)
Polyamide 6 [Nylon 6]	$-\!\left[NH-CH_2-CH_2-CH_2-CH_2-CH_2-CO\right]_n\!-$	4	230
Polyamide 6/6 [Nylon 6.6]	$-\!\left[NH-CH_2-CH_2-CH_2-CH_2-CH_2-CH_2-NH-CO-CH_2-CH_2-CH_2-CH_2-CO\right]_n\!-$	5	260
Polyethylene terephthalate [Polyester]		18	260
Polyethylenene naphthalate [PEN]		35	275
Vectran		80	330
Poly(*m*-phenylenediamine-isophthalamide) Nomex		17	400
Poly-*p*-phenylene/3,4-diphenylether terephthalamide [Technora]		70	500
Poly(*p*-phenylene terephthalamide) [Kevlar]		135	550
Poly(*p*-phenylene benzobisoxqazole) PBO [Zylon]		280	650
Poly (2,6-dimidazo(4.5-b:4′.5-e) pyridinylene-1,4 (2,5-dihydroxy) phenylene) (PIPD) M5		330	650

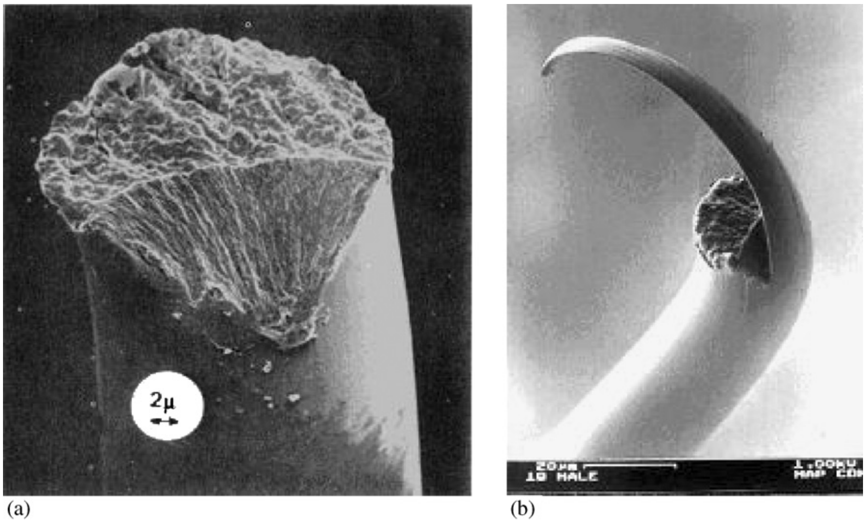


Figure 1
Fracture morphologies. (a) A polyamide 6/6 fiber broken in tension and (b) a polyamide 6/6 fiber broken in fatigue.

raising the cyclic loading pattern can stop fatigue failure.

Higher melting points or decomposition temperatures and much higher Young's moduli have been achieved by including more aromatic rings in the molecular main chain and by aligning these very rigid molecules parallel to the fiber (Table 1). These types of polymers are made by the technology of liquid crystal spinning, as illustrated in Fig. 2. Overall, within the melt or solution, the molecular alignment is random; however, orientation is achieved not by drawing but through the shearing of the solution when the solution passes through the spinneret. Molecular alignment in these fibers occurs because of the very stiff molecules that are aligned preferentially parallel to the fiber axis by the shear forces in the spinneret and held together by hydrogen or van der Waals' bonds. The straighter the molecule the higher the modulus achieved. These liquid crystal polymers do not usually lend themselves to melt spinning and are usually spun from a solution although the Vectran fiber, produced since the 1970s by Hoechst-Celenese, is melt spun.

Vectran is a liquid crystal polyester-based fiber, which, as can be seen from Table 1, has a molecular structure that is not perfectly straight. It is produced by the acetylation polymerization of *p*-hydroxybenzoic acid and 6-hydroxy-2-naphthoic. The good alignment of its molecular structure gives higher rigidity than that found with other polyester fibers but the form of the molecule determines its ultimate rigidity. The fiber is the only melt spun liquid crystal fiber, which is available and finds uses in applications such as sail cloth because of its rigidity and resistance in bending.

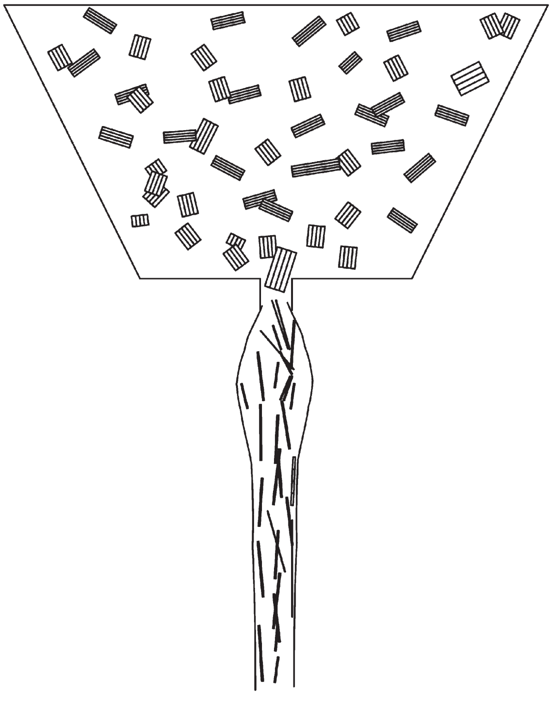


Figure 2
Spinning from a liquid crystal solution.

The aramid family of fibers are made up of aromatic PAs. Nomex was the first fiber of this family, produced in the 1960s by Du Pont. Nomex is a *m*-aramid, which accounts for an elastic modulus little different from high-performance PET but with a higher temperature resistance. Nomex is widely used for clothing for potentially hazardous conditions where fire is a danger. It is also widely used in paper form as a honeycomb structure in the aeronautical industry. *p*-Aramids can possess remarkably high Young's moduli, more than 20 times that of conventional PA fibers and twice that of glass, as shown in Table 2. This group of fibers, first introduced as the Kevlar fiber from Du Pont, has been available since 1972. These fibers are made from a mesophase or liquid crystal solution of poly-(*p*-phenylene terephthalamide) (PPTA) dissolved in concentrated sulfuric acid. If an appropriate concentration is made, the solution consists of locally organized PPTA molecules, which align with one another. The structure of aramid fibers is very well ordered; however, they are highly anisotropic, which results in them being weak in all directions except parallel to the fiber axis. As a consequence failure of these fibers is nearly always highly fibrillar.

Under tension, Kevlar fibers are 5 times as strong as steel in air for the same weight. Under tension, the fibers show little tendency to deform plastically, but when they are bent, the side experiencing compression undergoes plastic deformation, which allows them to accommodate the imposed strain. This results in the fiber being difficult to cut as well as giving great tenacity and resistance to impact loading. Their highly anisotropic behavior means that the fibers are not used in primary loading structures subjected to compressive forces. Aramid fibers, other than Kevlar, that are commercially available are the Twaron fiber developed by Akzo in Holland, now produced by Teijin Twaron, also in Holland and the Technora fiber from Teijin in Japan. Technora is derived from poly-*p*-phenylene/3,4′-diphenylether terephthalamide, which, as can be seen from Table 1, is not a straight molecule and as a consequence the fiber has an intermediate modulus of ~65 GPa.

In 1998 the Toyobo Company began the commercial production of the Zylon fiber. Its chemical composition is poly(*p*-phenylene benzobisoxazole) (PBO). The fiber has remarkable properties in tension, as can be seen from Table 2. It has a Young's modulus, which is 40% higher than that of steel, with the density of an engineering plastic. Its Young's modulus is twice that of Kevlar, since unlike Kevlar, the PBO molecule, shown in Table 1, is completely straight thus making the most of the covalent bonds in its structure. PBO fibers, however, suffer from poor compressive strength as molecular interchain cohesion is assured by van der Waals' bonds. As can be seen from Fig. 3, the fiber fibrillates on failure due to its inherent anisotropy.

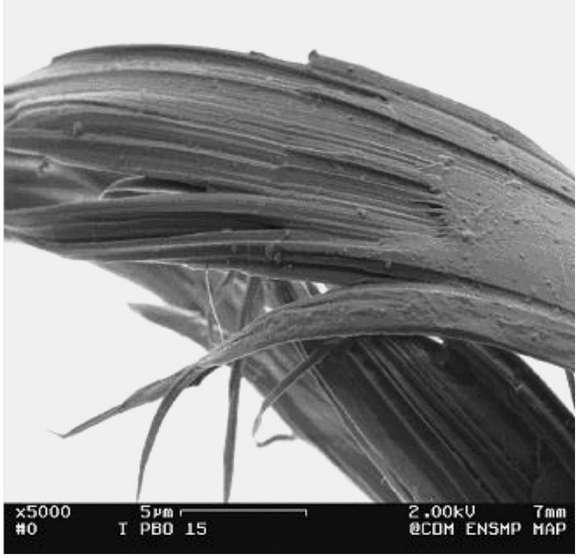


Figure 3
Detail of the splitting of a Zylon PBO fiber.

Table 2
Comparison of organic fiber properties.

Fiber	Diameter (µm)	Specific gravity	Young's modulus E (GPa)	Strength σ (GPa)	Strain to failure ε (%)	Specific modulus E/ρ (GPa g^{-1} cm^3)	Specific strength σ/ρ (GPa g^{-1} cm^3)
Polyamide 66	30	1.14	<5	1	20	4	0.9
Polyester	20	1.38	15	1	15	10	0.7
Pentex	28	1.4	35	1.1	6	25	0.8
Vectran	20	1.4	65	3.8	3.3	46	2
Nomex	15	1.4	17	0.7	20	12	0.5
Technora	12	1.39	70	3	4.4	50	2.2
Kevlar 49	12	1.45	135	3	4.5	93	2.1
Zylon	12	1.56	280	5.8	2.5	180	3.7
Polyethylene	38	0.96	117	3	3.5	120	3.1

Magellan International in the US has taken over the development of a fiber created by the researchers at Akzo, which at present is called M5. The fiber is said to possess a Young's modulus of 330 GPa, but, above all, superior compressive properties, such as molecular interchain cohesion, are ensured by hydrogen bonds. This leads to a slightly higher specific gravity of 1.7. Compressive strength is said to be 4 times higher than that of the PBO fibers and nearly 3 times higher than Kevlar fibers.

The production of aramid fibers involves complex and expensive chemistry; however, the high-performance properties of the fibers are a result of the molecules being aligned parallel to the fiber's axis. Simpler polymers could, if all their molecules were aligned parallel to one another, give materials with the characteristics of high-performance materials. Such fibers can be made using a sol-gel, in which polyethylene is dissolved in a solvent to make a very dilute solution, which is then spun. Such fibers were first produced in Holland by DSM, under the name Dyneema, and then by Allied (now Honeywell) in the US under the name Spectra. The fibers consist of polyethylene molecules, which are aligned parallel to the fiber axis with molecular interchain cohesion assured by van der Waals' forces. They have properties in tension rivaling those of the aramid fibers with a lower specific gravity (0.97). However their properties in compression are poor. These fibers are limited in temperature to a maximum of 120 °C as they melt around 150 °C and suffer from creep.

1.2 Glass Fibers

In 1931 two American firms, which were later to fuse to make the Owens Corning Fiberglas Corporation, developed a commercial method of spinning glass filaments from the melt through spinnerets. The melting point of the most widely used glass is around 1200 °C and the spinnerets are made from platinum. The holes in the spinnerets have diameters of 1–2 mm and drawing the filaments produces fibers having diameters usually between 5 and 15 μm. The spinnerets usually contain several hundreds of holes so that a strand of glass fibers is produced. Several types of glass exist but all are based on silica (SiO_2), which is combined with other elements to create specialty glasses. The compositions and properties of the most common types of glass fibers are shown in Table 3. The most widely used glass for fiber-reinforced composites is E-glass. S-glass, which contains a greater amount of alumina, is used to make glass fibers with superior mechanical properties whereas type C has a much reduced alumina content and resists an acid environment. Type D, which has no alumina, but a high content of boron oxide is used for its dielectric properties.

The strength of glass fibers is usually much higher than that of bulk glass as they contain far fewer weakening flaws. However to prevent damage by abrasion, they are coated with a protective size which also holds the strand together, acts as a bonding agent with resins and sometimes as an antistatic agent. The structure of glass fibers is amorphous with ionic bonds linking the atomic structure. There appears to be no appreciable residual stress within the fiber and the structure is isotropic, which means that the Young's modulus of glass in fiber form is the same as in bulk form. Glass fibers are cheaper than most other relatively high-modulus fibers and do not require very specialized machines or techniques to handle them. Their elastic modulus is, however, low compared to many other fibers and the specific gravity of glass is relatively high. Glass fibers are most commonly used to reinforce resins that limit their use to ⩽300 °C (usually ⩽120 °C). Above this range densification of the glass occurs and the fibers become increasingly susceptible to slow crack growth from surface defects. Strength increases with decreasing temperature.

Table 3
Compositions (wt.%), density, and mechanical properties of various glasses used in glass fiber production.

Glass type	E	S	R	C	D
SiO_2	54	65	60	65	74
Al_2O_3	15	25	25	4	
CaO	18		9	14	0.2
MgO	4	10	6	3	0.2
B_2O_3	8			5.5	23
F	0.3				
Fe_2O_3	0.3				
TiO_2					0.1
Na_2O				8	1.2
K_2O	0.4			0.5	1.3
Density	2.54	2.49	2.49	2.49	2.16
Strength (20 °C) GPa	3.5	4.65	4.65	2.8	2.45
Elastic modulus (20 °C) GPa	73.5	86.5	86.5	70	52.5
Failure strain (20 °C) %	4.5	5.3	5.3	4.0	4.5

1.3 Deposited Fibers

Boron, the fifth element in the periodic table, is the lightest element with which it is practical to make fibers and they were first produced in the US at the beginning of the 1960s. Boron fibers are made by chemical vapor deposition, in which a mixture of boron trichloride (BCl_3) and hydrogen (H_2) at around 1000 °C passes over a continuous tungsten filament heated by electric current. Boron is

Table 4
Properties of ceramic fibers produced by chemical vapor deposition.

Fiber	Manufacturer	Trade mark	Composition	Diameter (μm)	Density ($g\,cm^{-3}$)	Strength (GPa)	Strain to failure (%)	Young's modulus (GPa)
SiC	Specialty materials (formerly Textron)	SCS-6	SiC on carbon	140	2.7–3.3	3.4–4.0	0.8–1	427
	DERA	Sigma	SiC on tungsten	100	3.4	3.4–4.1	0.8	400–410
B	Speciality materials (formerly Textron)		Boron on tungsten	100–140	2.57	3–6	1	380–400

deposited on the tungsten filament, which has a diameter of 12 μm.

The fibers have Young's moduli exceeding 400 GPa as shown in Table 4 and very high strength both in tension and in compression. This latter quality is enhanced by their large diameter of 140 μm. A cross section of a boron fiber reveals its tungsten core and boron mantle. Boron fibers have proved to be a good reinforcement for light metal alloys particularly when the fiber was protected by a coating of SiC or B_4C. However, their high cost of production has limited the wide use of this type of fiber.

Silicon carbide fibers are also produced by CVD using various chlorosilanes, such as CH_3-Si-Cl_3, which gives 3HCl and SiC, which is deposited onto a tungsten or carbon core. They are of primary interest for the reinforcement of titanium and intermetallic matrices. Fibers with diameters of between 100 and 140 μm are made. These fibers, until recently produced by Textron and now produced by Speciality Materials Inc., are known as SCS-*n*, where *n* is the number of microns of surface coating added onto the diameter to protect it during composite manufacture. SiC fibers have a Young's modulus $\geqslant$400 GPa, as shown in Table 4 and can be used to around 1000 °C. Figure 4 shows the cross section of an SiC fiber made by CVD showing the tungsten core.

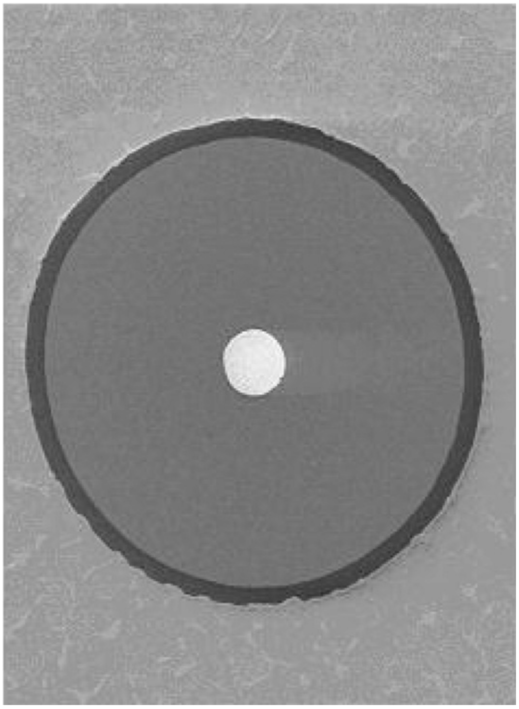


Figure 4
Cross-section of an SiC fiber made by CVD showing the tungsten core.

1.4 Carbon Fibers

Carbon is the sixth lightest element and the carbon–carbon covalent bond is the strongest in nature. Many fibers can be converted into carbon fibers, the basic requirement being that the precursor fiber carbonizes rather than melts when heated. The carbon fibers developed in the US in the 1950s and early 1960s used fibers regenerated from cellulose. This proved a slow process, as the carbon yield from cellulose is only 24% and mechanical properties were not high. However, such fibers have high thermal conduction properties and are still used in carbon–carbon heat shields and brake pads.

Most carbon fibers though are made by a process developed first in Great Britain and then in Japan in the late 1960s based on the conversion of a modified form of acrylic textile fiber. Polyacrylonitrile (PAN) has a carbon yield of 49% and the properties of carbon fibers made by this method depend on the temperature of pyrolysis. The PAN fibers, held in tension, are first heated to about 250 °C in air, which cross links the structure and makes it infusible. Further heating in an inert atmosphere to above 1000 °C eliminates most of the elements other than carbon and around 1500 °C a fully carbon fiber is achieved. At this temperature the strength of carbon fibers reaches a maximum. Above this temperature, the basic carbon structural units increase in size, which results in a fall in strength whereas the Young's modulus increases continuously with increasing temperature up to around 3000 °C. The structure, which results from the pyrolysis of PAN, is highly anisotropic, with the basic structural units formed by the carbon atom groups aligned parallel to the fiber axis. There is complete rotational disorder in the radial direction and the relatively poor stacking of the carbon atoms means that a graphite structure is never achieved. The structure contains pores that account for the density of the fibers being less than that of fully dense carbon. Table 5 shows some typical properties of carbon fibers; however, there exists a considerable range of properties for these fibers. Figure 5 shows a PAN-based carbon fiber.

Carbon fibers made from pitch-based precursors were developed in the US in the early 1970s and in Japan in the 1980s. The high carbon yield of pitch, which approaches 90%, makes it an attractive and cheap source for making carbon fiber precursors. However, cost is increased by the purification processes by which the pitch is converted into a mesophase or liquid crystal solution, which is then spun giving precursor fibers with aligned microstructures. From this point the fabrication route is the same as that for the PAN-based fibers. Unlike PAN-based fibers, there is no peak in strength at 1500 °C. Pitch-based precursors heated to around 2300 °C give fibers with Young's moduli as high as those obtained with PAN-based fibers at 2900 °C and heating to these higher temperatures gives even greater stiffness of up to four and a half times that of steel. It is therefore more economical to produce high-modulus carbon fibers from pitch than from PAN. The properties are due to a less disordered, more graphitic microstructure of the pitch-based fibers, which leads to lower compressive strength and increased costs in producing high-strength carbon fibers.

In the absence of oxygen, carbon fibers are the ideal reinforcement and can be used above 3000 °C. However, the fibers suffer from oxidation from around 400 °C so that they can be used for short exposure times, such as in carbon–carbon composite rocket nozzles or heat shields for atmospheric re-entry vehicles, but not for long-term applications above this temperature.

1.5 Oxide Fibers

The oxide of aluminum, alumina, is a hard brittle material with excellent refractory properties up to 1600 °C. The atomic structure of this material is assured by a mixture of covalent and ionic bonds. Fibers for structural applications based on forms of alumina were developed first in the UK and in the USA in the 1970s. Alumina exists in a range of phases of which the most stable and crystalline form, to which every other phase is converted around 1200 °C, is alpha alumina which has a Young's modulus of around 400 GPa. Several manufacturers have produced continuous α-alumina fibers as can be seen

Table 5
Typical characteristics of carbon fibers.

Fiber type	Manufacturer	Trade mark	Diameter (μm)	Density ($g\,cm^{-3}$)	Strength (GPa)	Strain to failure (%)	Young's modulus (GPa)
Ex PAN							
High strength	Toray	T400H	7	1.80	4.4	1.8	235
High strength	Toray	T1000	5	1.82	7.1	2.4	294
High modulus	Toray	M46J	7	1.84	4.2	1.0	436
High modulus	Toray	M60J	5	1.94	3.9	0.7	588
Ex pitch							
Oil-derived pitch	Nippon graphite fiber	Granoc CN-35	11	2.10	3.6	1.0	350
Oil-derived pitch (high modulus)	Nippon graphite fiber	Granoc YS-95A	11	2.16	3.5	0.4	920
Coal tar pitch	Mitsubishi chemicals	Dialead K1352U	10	2.12	3.6	0.58	620
Coal tar pitch	Mitsubishi chemicals	Dialead K13B2U	10	2.16	3.9	0.48	830

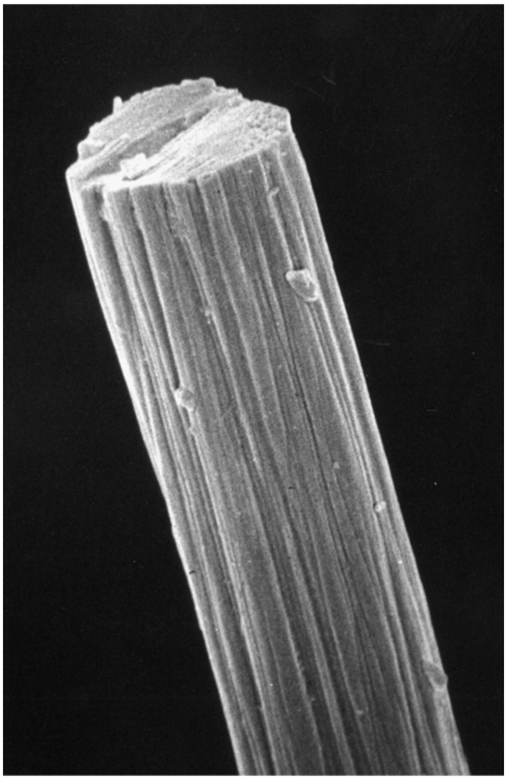


Figure 5
A PAN-based carbon fiber.

from Table 6. These fibers are made from slurries consisting of α-alumina particles, in an aqueous solution of aluminum salts which is spun and sintered. The control of grain growth and porosity in the production of α-alumina fibers is difficult so that the different manufacturing routes adopted have produced fibers with different Young's moduli. The Almax fiber, for example, contains considerable porosity, which accounts for its lower modulus than that of Nextel 610.

The difficulties in producing pure α-alumina fibers can be overcome by the inclusion of silica in the structure and the development of intermediate alumina phases. The Young's moduli, around 200 GPa, of these fibers are lower compared to that of pure α-alumina and such fibers are produced at a lower cost. This, added to easier handling due to their lower stiffness, makes them attractive for thermal insulation applications, up to at least 1500 °C and as reinforcements for aluminum alloys in the temperature range of 300–350 °C. α-Alumina fibers and those of other phases containing silica resist oxidation at high temperature but the high diffusion rates of the elements in their structures above 1000 °C means that they creep and also lose strength around this temperature.

A series of fibers based on a mullite composition consisting of a ratio of 3 moles of alumina to 2 of silica has been developed by 3 M. Mullite has a complex crystallographic structure, which is highly resistant to creep at high temperatures. The Nextel 720 fiber from 3 M consists of both α-alumina and mullite phases. The microstructure of the fiber is composed

Table 6
Properties and compositions of alumina-based fibers.

Fiber type	Manufacturer	Trade mark	Composition (wt.%)	Diameter (μm)	Density ($g\,cm^{-3}$)	Strength (GPa)	Strain to failure (%)	Young's modulus (GPa)
α-Al_2O_3 fibres	Mitsui mining	Almax	99.9% Al_2O_3	10	3.6	1.02	0.3	344
	3M	610	99% Al_2O_3, 0.2%–0.3 SiO_2, 0.4–0.7 Fe_2O_3	10–12	3.75	1.9	0.5	370
Alumina silica fibres	Saffil	Saffil	95% Al_2O_3, 5% SiO_2	1–5	3.2	2	0.67	300
	Sum itomo	Altex	85% Al_2O_3, 15% SiO_2	15	3.2	1.8	0.8	210
	3M	312	62% Al_2O_3, 24% SiO_2, 14% B_2O_3	10–12 or 8–9	2.7	1.7	1.12	152
	3M	440	70% Al_2O_3, 28% SiO_2, 2% B_2O_3	10–12	3.05	2.1	1.11	190
	3M	720	85% Al_2O_3, 15% SiO_2	12	3.4	2.1	0.81	260

Table 7
Properties and compositions of silicon-based fibers.

Fiber type	Manufacturer	Trade mark	Composition (wt.%)	Diameter (μm)	Density (g cm^{-3})	Strength (GPa)	Strain to failure (%)	Young's modulus (GPa)
Si-C based	Nippon carbide	Nicalon NLM 202	56.6% Si 31.7% C 11.7% O	14	2.55	2.0	1.05	190
	Ube industries	Tyranno Lox-M	54.0% Si 31.6% C 12.4% O 2.0% Ti	8.5	2.37	2.5	1.4	180
	Nippon carbide	Hi-Nicalon	62.4% Si 37.1% C 0.5% O	14	2.74	2.6	1.0	280
SiC	Nippon carbon	Hi-Nicalon S	SiC + O + C	13	3.0	2.5	0.65	375
	Ube industries	Tyranno SA	SiC + C + O + Al	10	3.0	2.5	0.75	330
	Dow corning	Sylrami c	SiC + TiB2 + C + O	10	3.1	3.0	0.75	390

of mosaic grains of 0.5 μm with wavy contours, consisting of several slightly mutually misoriented grains and elongated grains. Its Young's modulus is 260 GPa. At high temperatures, above 1000 °C and up to 1500 °C, the microstructure of the Nextel 720 fiber leads to a dramatic reduction in creep rate compared to pure alumina fibers. However, the fiber surface is particularly reactive to alkaline contamination above 1100 °C, which leads to the development of a granular surface and a subsequent fall of strength occurs.

1.6 SiC-based Fibers

Fibers, based on silicon carbide, became available from Japanese producers in the early 1980s. They are made by the pyrolysis of organo-silicon precursor filaments based on polycarbosilane, which consists of cycles of six atoms arranged in a similar manner to the diamond structure of β-SiC. The early Nicalon fibers made by Nippon Carbon and the Tyranno fibers from Ube Industries were prepared by cross-linking the precursor fibers by heating in air, which introduced oxygen into the structure. The Tyranno fibers already contained some oxygen as it was introduced with a small amount of titanium into the precursor to control spinning. The presence of the oxygen resulted in the SiC grains, about 2 nm in size, which were developed during pyrolysis, being surrounded by an amorphous Si–C–O phase and the fibers had Young's moduli of around 200 GPa. This intergranular phase caused the fibers to creep above 1000 °C and its degradation above 1200 °C limited their use. Cross-linking of the precursors by electron irradiation eliminated the need for introducing additional oxygen into the precursors. This process gave rise to the Hi–Nicalon fiber, which had no significant oxygen-rich amorphous phase, but considerable excess free carbon. The grain size of the Hi–Nicalon fiber is around 10 nm and its Young's modulus is in the range of 265–290 GPa. The fiber is more stable than its predecessors at temperatures above 1000 °C although grain growth occurs. Both the Japanese manufacturers and Dow Corning have produced a near stoichiometric SiC fiber by extending the above technology and heating the fibers to temperatures of around 1800 °C. The fibers differ in detail, which leads to different ultimate properties, as can be seen in Table 7, but all three have grain sizes greater than 50 nm and up to 200 nm. This leads to Young's moduli of up to 400 GPa. These fibers retain their properties in short-term tests to around 1400 °C.

2. Concluding Remarks

A wide range of very different high-performance fibers have been developed. Many have low densities making them attractive engineering materials, either embedded in a matrix to form a composite, or by themselves as ropes or other structures. These materials, which are still being improved, are destined to find a wide range of uses some of which would be impossible without the extraordinary properties of high-performance fibers.

Bibliography

Bunsell A R 1988 *Fiber Reinforcements for Composite Materials*. Elsevier, Amsterdam

Bunsell A R, Berger M-H 1999 *Fine Ceramic Fibers*. Dekker, NY

Chawla K K 1998 *Fibrous Materials*. Cambridge University Press, Cambridge

Hearle J W S 2001 *High-performance Fibres*. Woodhead Publishing Ltd., Cambridge

Moncrieff R W 1979 *Man-made Fibers*, 6th edn. Butterworth, Guildford

Kostikov V I 1995 *Fiber Science and Technology*. Chapman and Hall, London

A. R. Bunsell
Centre des Matériaux, Evry,
France

I

Inorganic and Inorganic–Organic Aerogels

Aerogels are highly porous solid materials with very low densities and high specific surface areas. Their structure consists of a filigrane solid network with open, cylindrical, branched mesopores. This results in interesting physical properties, such as extremely low thermal conductivity and low sound velocity, combined with high optical transparency. Aerogels can be obtained as monoliths, granulates, films, or powders. There are several review articles (Ayen and Iacobucci 1988, Fricke and Emmerling 1992, Gesser and Goswami 1989, Heinrich *et al.* 1995, Hüsing and Schubert 1998) and symposium proceedings on aerogels (Fricke 1986, 1992, Pekala and Hrubesch 1995, Phalippou and Vacher 1998).

1. Network Formation

A gel consists of a sponge-like solid network, the pores of which are filled with a gas or a liquid. When gels are prepared from solutions of molecular precursors (through the sol stage), the pores are filled with the solvent ("wet" gels). Aerogels are obtained when the pore liquid in the wet gel is replaced by air without decisively altering the highly porous network structure or the volume of the gel body. In contrast, xerogels (see *Xerogels**) are formed by conventional drying of the wet gels with concomitant large shrinkage (and mostly destruction) of the gel body.

The procedures for the preparation of xerogels and aerogels are the same up to the wet gel stage, but the drying process is different. However, because the properties of aerogels very critically depending on the structure of the gel network, the chemical parameters governing the network formation have to be controlled more deliberately than required for most xerogels.

Inorganic or inorganic–organic gels are prepared by sol–gel processing (Brinker and Scherer 1990) from purely inorganic precursors (hydrolyzable metal salts or oxide sols) or from metal or semi-metal alkoxides and their organically substituted derivatives (see *Sol–Gel Reactions: Theory*#). The network and thus the pore structure of gels is formed by the aggregation of small primary colloidal or polymeric (sol) particles with typical diameters of 1–3 nm. In colloidal gels, dense colloidal primary particles are interconnected like a string of pearls. In polymeric gels, linear or branched polymer chains are formed by condensation of small clusters (the oligomeric species formed during sol–gel processing). There are several theoretical models for the aggregation of the primary particles to three-dimensional networks (Brinker and Scherer 1990, Schaefer 1989).

Aerogel networks are usually formed from colloidal particles. However, each aerogel has its own structural characteristics because the microstructure strongly depends on the preparation conditions.

1.1 Silica-based Aerogels

The term "SiO_2 gel" (correspondingly for other gels) is used to characterize the type of inorganic skeleton. "SiO_2 gels" and the corresponding aerogels often have the composition $SiO_x(OH)_y$ or, if they are prepared from alkoxides, $SiO_x(OH)_y(OR)_z$, where the values of y and z can be rather high.

Most laboratory preparations of silica aerogels use tetraalkoxysilanes ($Si(OR)_4$) as the silica source. The less expensive aqueous sodium silicate solutions (water glass) are used for larger-scale technical applications. The sodium silicate solution is ion-exchanged and the resulting silicic acid solution gelled by lowering the pH.

The kind of inorganic network formed by sol–gel processing depends to a large extent on the relative rates of the hydrolysis and condensation reactions. Hydrolysis is the rate-determining step under basic conditions, where reaction at central silicon atoms of an oligomeric unit is favored. The network of the resulting colloidal gels consists of big particles and large pores. The clusters mainly grow by the condensation of monomers, because condensation of clusters is relatively unfavorable. Acidic conditions ($pH = 2–5$) favor hydrolysis, and condensation is the rate-determining step. A great number of monomers or small oligomers with reactive Si–OH groups is simultaneously formed. Reactions at terminal silicon atoms are favored. This results in polymer-like networks with small pores.

Complete removal of the pore liquid from the smaller pores in polymeric gels is more difficult, resulting in a larger shrinkage during drying. For this reason, silica aerogels are usually prepared by base-catalyzed reaction of $Si(OR)_4$.

A modification of this procedure is to prehydrolyze $Si(OR)_4$ with a small amount of water under acidic conditions. In a second step, a defined amount of aqueous acid or base is added. Base catalysis in the second step results in a stiffening of the network that stabilizes the gels. This two-step procedure allows a more deliberate control of the microstructure of the SiO_2 gels.

The macroscopic properties of SiO_2 aerogels differ widely, due to structural differences. Silica aerogels with bulk densities as low as $0.003\,g\,cm^{-3}$ (99.8% porosity) have been prepared. Typical values are in the range $0.1–0.2\,g\,cm^{-3}$ (85–90% porosity). Skeletal densities are in the range $1.7–2.1\,g\,cm^{-3}$, mean pore

---

* Cross references marked by an asterisk are included in this volume.

# Cross references marked by a hash can be found by consulting the Encyclopedia of Materials: Science and Technology.

diameters are in the range 20–150 nm, and specific surface areas are in the range 100–1600 $m^2 g^{-1}$. Silica aerogels are always amorphous.

The spectrum of properties can be widened by modification of silica aerogels with organic entities. For example, the hydrophobicity and the elastic properties of SiO_2 aerogels are improved by incorporation of organic groups, and new applications can be envisioned by the integration of functional organic moieties. Any modification should retain the gel network and the pore structure, because the interesting physical properties of aerogels result from these. Possibilities of postsynthesis doping or modification of aerogels with organic compounds are limited. Therefore, the organic groups have to be incorporated during sol–gel processing. Embedding molecules or polymers in gels is achieved by dissolving them in the precursor solution. The gel matrix is formed around them and traps them. The doped wet gels can be converted to aerogels. However, the probability is very high that the organic groups are leached out during the drying process.

A more general route for the organic modification is to use hydrolyzable and condensable precursors for sol–gel processing in which the organic group is covalently attached to the silicon atom and will therefore be bonded to the inorganic network. Silica aerogels modified by non-functional or functional organic groups are prepared by sol–gel processing of $R'Si(OR)_3/Si(OR)_4$ mixtures followed by supercritical drying of the wet gels. The process can be controlled in such a way that the organic groups cover the inner surface of the gel network without influencing its basic structure.

1.2 Metal Oxide Aerogels

The principles for network formation of non-silicate inorganic gels are the same as for SiO_2 gels. Aqueous salt solutions or metal alkoxides are employed for sol–gel processing. Metal alkoxides are much more reactive towards water than alkoxysilanes. This is due to the lower electronegativity and higher Lewis acidity as well as the possibility of increasing the coordination number. Whereas the reactivity of alkoxysilanes has to be promoted by catalysts, the reaction rates of metal alkoxides must be moderated to obtain gels instead of precipitates. The most common method is the substitution of part of the alkoxy groups by bidentate anionic ligands, such as carboxylates, β-diketonates, etc.

Unlike silica aerogels, their non-silicate counterparts can also form crystalline primary particles. As a general rule, the crystalline portion is favored by a large excess of water in the hydrolysis reaction. For example, TiO_2 aerogels can be prepared totally amorphous or with a network of mesoparticles (about 50 nm in diameter) formed by aggregation of crystalline anatase nanoparticles of about 5 nm diameter. Amorphous and partially crystalline ZrO_2 aerogels are found to have a branched, polymer-like fractal structure consisting of small primary particles with average diameters of 5.2 nm (partially crystalline) or 2.5 nm (amorphous).

One-component oxide aerogels include titania, alumina, and zirconia (mainly prepared from the propoxides and butoxides), V_2O_5 (prepared from $VO(OR)_3$), Cr_2O_3 (prepared from CrX_3, $X = NO_3$, Cl, or OAc), Fe_2O_3 (prepared from $FeCl_3$ or $Fe(acac)_3$), MoO_3 (prepared from $MoO_2(acac)_2$), and Nb_2O_5 (prepared from $Nb(OEt)_5$).

Mixed-metal, binary, or ternary aerogels can be prepared from the same types of precursors. Typical examples are TiO_2/SiO_2, mullite, Fe_2O_3/SiO_2, Fe_2O_3/Al_2O_3, V_2O_5/MgO, PbO/Al_2O_3, PbO/ZrO_2, BaO/Al_2O_3, Li_2O/B_2O_3, $PbTiO_3$, $NiO/SiO_2/MgO$, and cordierite. Many of these were prepared for catalytic applications (Schneider and Baiker 1995, 1997; Pajonk 1991, 1997).

When mixtures of two or more precursors with very dissimilar hydrolysis and condensation rates are reacted, the microstructure of the product may become rather heterogeneous, and phase separation may even occur. The faster reacting compound may form sol particles, which are coated by the slower reacting component. This effect can be avoided by moderating the reaction rate of the faster reacting component by the already discussed bidentate ligands. Alternatively, the slower reacting precursor can be prehydrolyzed and the faster reacting component added afterwards.

In mixed-oxide aerogels, the SiO_2 and Al_2O_3 portions are always amorphous. However, crystalline aluminates are observed (e.g., $NiAl_2O_4$). The structural data and properties of non-silicate aerogels depend very strongly on the particular system and on the preparation conditions. Typical densities are in the range 0.1–0.3 $g\,cm^{-3}$, and typical porosities in the range 80–95%. Pore radii range from 1 to 25 nm, specific surface areas from 80 to 800 $m^2 g^{-1}$, and primary particle diameters from 2 to 50 nm.

1.3 Metal–Aerogel Composites

A second metal component in the gels may not necessarily yield a binary oxide aerogel. The use of alcohols for supercritical drying can, under the applied conditions, result in the reduction of oxides to metals. Metal particles can also be generated by subsequent reduction of the aerogels with hydrogen at high temperatures or by mixing hydrogen into the autoclave during supercritical drying.

Postsynthesis doping of aerogels with metals is achieved by impregnation with the alcoholic solution of a suitable metal salt. The solvent has then to be removed again by another supercritical drying step in order to prevent destruction of the aerogel network (two-step process). Composites prepared by this approach include Cu/SiO_2, Cu/MgO, Pt/MoO_3, Pd/Al_2O_3, and Pd/CeO_2. Alternatively, suitable metal compounds can be mixed into the precursor solution, which results in the incorporation of the metal compounds in the gel during sol–gel processing. Examples of metal-containing aerogels prepared by this approach include Pt/TiO_2, Pt/Al_2O_3, Pd/Al_2O_3, V/SiO_2, and Cu/SiO_2.

2. Drying Methods

The evaporation of the liquid from a wet gel is a very complex process. When wet gels are conventionally dried, the slower shrinkage of the network in the interior of the gel body results in a pressure gradient. Furthermore, the meniscus of the liquid drops faster in larger pores. The walls between pores of different size are therefore subjected to uneven stress and cracking (Brinker and Scherer 1990). The capillary forces exerted by the meniscus of the pore liquid and the pressure gradient exerted by the large shrinkage of the network are the main reasons for the collapse of the filigrane structure upon conventional drying. For the production of aerogels, methods must be applied which conserve the pore structure of the wet gels.

2.1 Supercritical Drying

In this procedure, the pore liquid is put into the supercritical state, where the liquid/gas interface disappears. The wet gel is placed in an autoclave and both the temperature and the pressure are adjusted to values above the critical point of the corresponding solvent (p_c, T_c). The fluid is then slowly vented at constant temperature, while the pressure drops. When ambient pressure is reached, the vessel is cooled to room temperature. Often the vessel is prepressurized by nitrogen to avoid evaporation of the solvent. The gas/liquid phase boundary must not be crossed during the drying process.

The pore liquid itself, which is usually an alcohol or acetone from the wet gel preparation, can be put into the supercritical state; about 250 °C and 7000 kPa are required. Problems arise from the combination of high temperatures and high pressures, and the flammability of these solvents. Furthermore, rearrangement reactions in the gel network are highly probable because of the high temperatures. As a consequence, the resulting aerogels have a lower specific surface area, a more narrow pore radii distribution (elimination of microporosity), and a stiffer network. Organic groups may be destroyed during drying. Phase separation and loss of stoichiometry or crystallization may occur in multicomponent gels.

An alternative is the use of liquid CO_2. This has the advantage of a low critical temperature ($T_c = 31.1\,°C$) at a moderate critical pressure ($p_c = 7392\,kPa$). However, a time-consuming solvent exchange is necessary before the actual drying step. Another requirement is the miscibility of the pore liquid with carbon dioxide. Structural changes only take place to a minor extent with CO_2 as the supercritical fluid.

2.2 Freeze Drying

Another possible way to avoid the gas/liquid interface during drying is by freeze drying. The pore liquid is frozen and then sublimed *in vacuo*. However, the aging period has to be extended to stabilize the gel network, the solvent must be exchanged by another with a low thermal expansion coefficient and a high pressure of sublimation, and low freezing temperatures have to be achieved by addition of salts. Another disadvantage is that the network may be destroyed by crystallization of the solvent in the pores. Therefore, only powders are obtained.

2.3 Ambient Pressure Drying

Drying of the gels at ambient pressure requires strengthening of the network by exact control of the aging conditions to avoid its collapse.

One approach involves a series of solvent exchange processes and a modification of the inner surface. The water/alcohol mixture in the pores of the gel is first exchanged for a water-free solvent, and the Si–OH groups at the inner surface are silylated. The gel surface is thus hydrophobized and its reactivity reduced. In the first stage of the drying process the gel shrinks strongly, as expected. However, no irreversible narrowing of the pores by formation of Si–O–Si bonds is possible because of the silylation. Therefore, the gel expands to nearly its original size in the final stage of the drying process (“spring-back” effect). The network of the modified gels must be stable enough to tolerate a reversible shrinkage of up to 28% of its original volume.

In another strategy, the strength and stiffness of the network is increased by aging the wet gels in solutions of tetraalkoxysilanes in aqueous alcohols. The monomers condense in the smallest pores and at the particle necks, the microporosity of the gels is lost, and shrinkage during drying can be completely avoided. The gels can be dried at ambient pressure at temperatures of 20–180 °C due to the stiffer network.

With the ambient pressure drying technique for the preparation of aerogels, a simpler and therefore less expensive preparation method is available, which makes the industrial application of aerogels economically more interesting. Ambient pressure drying also facilitates the possibility of preparing aerogel films.

3. Postsynthesis Modification of Aerogels

Postsynthesis treatment with a liquid is difficult because wet gels are again formed. Doping with metal salts by impregnation has already been mentioned in Sect. 1.3.

Another, more general possibility for the postsynthesis modification of aerogels is their reaction with gaseous compounds. Organic compounds can be sublimed into the network as shown for a porphyrin derivative. Carbon was deposited on the inner surface of aerogels by decomposition of gaseous organic compounds to improve the thermal properties of aerogels.

The chemical vapor infiltration method has been employed to prepare Si/SiO_2 nanocomposites by reaction of SiH_4 or $HSiCl_3$, and for the deposition of iron or tungsten particles from ferrocene, $Fe(CO)_5$ or $W(CO)_6$. In an oxidizing atmosphere, amorphous iron oxide particles in SiO_2 aerogels were obtained which in subsequent reactions were transformed to Fe_3O_4 or Fe_9S_{10}-doped SiO_2 aerogels. WN or WS_2 particles in aerogels were formed by post-treatment of the tungsten-doped aerogels with ammonia or sulfur.

The coating of the inner aerogel surface with carbon structures can also be achieved by pyrolysis of organic groups covalently bonded to the gel network.

4. Properties

A big problem with unmodified SiO_2 aerogels, particularly for technical uses, is their long-term stability in humid atmospheres. Owing to the large number of silanol groups on their inner surface, adsorption and capillary condensation of water and eventually cracking of the gel body by the resulting capillary forces results. Hydrophobation of the aerogels is achieved by silylation reactions via the gas phase, or by treatment of the wet gel with a silylating reagent or by using organically modified precursors during sol–gel processing.

The optical properties of aerogels strongly depend on their chemical composition. SiO_2 aerogels and the derived inorganic–organic hybrid aerogels vary between transparent and translucent. They appear yellowish under light and bluish against a dark background due to Rayleigh scattering. Objects viewed through a silica aerogel tile appear slightly blurred. For most applications a high optical transmission is desirable. Improvements were, *inter alia*, achieved by optimizing the drying procedures and the synthesis conditions.

Although the rigidity of aerogels is low due to their structure and the brittleness of the SiO_2 particles, it is sufficient for most applications. The compressive strength of monolithic SiO_2 aerogels is between 0.15 and 0.30 $\mathrm{N\,mm^{-2}}$, depending on the density, with an elastic compression of about 2–4%. The rigidity in vacuum is distinctly higher. The tensile strength is about 0.020 $\mathrm{N\,mm^{-2}}$ and the modulus of elasticity 0.002–100 MPa.

One of the decisive parameters for mechanical properties is the connection of the primary particles of the gel network. Silica aerogels with covalently bonded organic groups have improved elastic properties because the organic groups at the surface of the particles impair stiffening of the particle necks by Ostwald ripening.

The sound velocities in SiO_2 aerogels of 100–300 $\mathrm{m\,s^{-1}}$ are among the lowest for inorganic solids due the particular network structure. SiO_2 aerogels have the lowest acoustic impedance of all solid materials ($Z = 104\text{–}105\,\mathrm{k\,g\,m^{-2}\,s^{-1}}$).

Aerogels have the lowest thermal conductivities of all solids (0.017–0.021 $\mathrm{W\,m^{-1}\,K^{-1}}$ at 300 k in air). The solid conductivity portion increases with increasing bulk density, whereas the reverse is true for gas and radiative transport. Heat transport via the gas phase can be avoided by evacuation of the system. For example, an evacuated SiO_2 aerogel at about 300 k shows a thermal conductivity of only 0.010 $\mathrm{W\,m^{-1}\,K^{-1}}$. Reduction of the radiative heat transport is important for high-temperature applications of aerogels, because the radiation maximum is then in a wavelength range in which SiO_2 does not sufficiently absorb (3–5 μm). This is achieved by doping the aerogels with infrared opacifiers, such as carbon soot or TiO_2.

5. Applications

SiO_2 aerogels are the best investigated aerogels with regard to their preparation, structure, and properties, and they can also be prepared at lower cost than other aerogels. There are already commercial applications. Non-silicate aerogels increasingly play an important role, particularly as novel catalysts.

SiO_2 aerogels are installed in several Cherenkov detectors worldwide. Their refraction indices (1.007–1.024) are in a range not covered by the previously used compressed gases or liquids. Aerogels as solid materials additionally allow easier construction of the detectors.

Aerogels belong to the best thermal insulation materials currently available. Additional advantages are their nonflammability and their transparency. However, since the production of large aerogel panels is still a big problem, conventional window systems cannot be equipped. The commonly prepared granulate, which is not transparent due to the roughness of its surface, can only be used for "daylight windows" (e.g., bathroom, staircase, or ceiling windows). In addition to transparent thermal insulation, applications for a variety of other insulation problems are conceivable, for example in cooling or heating systems, high-temperature batteries, insulation for piping, heat and cold storage appliances, automotive exhaust pipes, transport vehicles, and vessels. Aerogels can also be utilized for passive use of solar energy, for example for paneling of house walls or for coating solar-energy collectors. A layer of the transparent aerogel allows the penetration of the sun's radiation to the blackened wall, but not the escape of the heat generated there.

All the advantages of sol–gel processing for preparing catalysts or catalyst supports—in particular, the controllable dispersion of the active component and the possibility of deliberately influencing the microstructure—also apply for aerogels. Additionally, the very high and accessible inner surface, the large pore volumes, an excellent temperature stability, and the high selectivities in catalytic processes predestines them for these applications. Nearly all oxides relevant for catalytic applications can be prepared as aerogels, including binary and ternary compositions, mixtures of metal oxides, and metal particles on oxide carriers. The investigated systems are summarized in special reviews (Schneider and Baiker 1995, 1997, Pajonk 1991, 1997).

The high porosity and very large surface area of silica aerogels can be utilized for applications as absorbing and encapsulation media, and hydrogen fuel storage. Partially sintered aerogels that show a strengthened texture can be used for the storage, thickening, or transport of liquids, for example rocket fuels.

Aerogels can be sintered at low temperatures and can therefore be processed to extremely pure and totally homogeneous glasses. The inner surface area and the porosity decrease during sintering and the pore structure can be modified in a controlled way by partial sintering. This was used, for example, for producing gas filters with pores in the range of 20–100 nm. Doping of aerogels with dyes or rare earth elements allows the fabrication of glass for various applications including lasers and radioluminescent light sources.

Aerogel films can be exploited as optical coatings for solar cells because of their low refractive index, and can thus replace the currently used glass layers with high refractive indices. More light reaches the active surface because of the lower Fresnel scattering losses. These films have also been applied as coatings for fiber optics to improve the light collection and propagation efficiency.

Aerogels are excellent materials for low dielectric constant applications. While bulk aerogels can be used for microwave electronics and high-voltage insulators, thin aerogel films can compete with existing technologies in microelectronics, where the low dielectric constant of <2.0 plays a particularly important role. In addition, the dielectric constant is easily controlled because it is directly correlated with the porosity and thus the density of the materials.

Some metal oxide aerogels exhibit superconducting, thermoelectric, or piezoelectric behavior.

Aerogels are efficient sound insulation materials. They have very low acoustic impedances and can therefore be used as $\lambda/4$ layers for matching the high acoustic impedance of piezoceramics to the low impedance of air.

SiO_2 aerogels are currently one of the standard media for collecting cosmic dust. They are fixed on the outside of spacecraft. Aerogels have the advantage that they are extremely light. The unique mesostructure enables a "soft landing" of very fast extraterrestrial particles hitting with a velocity of more than 3 km s^{-1}, and also allows their later investigation by optical methods.

6. Future Perspectives

The unique optical, thermal, acoustic, and mechanical properties of aerogels originate from the combination of a solid matrix (the chemical composition of which can be modified) and nanometer-sized pores filled with air. There is a direct connection between the chemistry of the sol–gel process and the structure of the gels on one hand, and between the structure and the properties of the aerogels on the other.

The most important area for the application of aerogels is thermal insulation, while important acoustic insulation applications are emerging. There is no doubt about the physical and ecological advantages (nontoxic, nonflammable, easily disposed) of SiO_2 aerogels compared with most other materials on the market.

Apart from special applications in which material costs only play a minor role, the rather high price of supercritically dried aerogels has prevented a broader range of applications. The new ambient pressure drying techniques will probably make the technical preparation much cheaper and will thus make aerogels more competitive. They will also allow the preparation of aerogels with standard laboratory equipment.

Bibliography

Ayen R J, Iacobucci P A 1988 Metal oxide aerogel preparation by supercritical extraction. *Rev. Chem. Eng.* **5**, 157–98

Brinker C J, Scherer G W 1990 *Sol–Gel Science*. Academic Press, New York

Fricke J 1986 (ed.) *Proc. 1st Int. Symp. Aerogels. Springer Proc. Phys.* **6**, Springer Berlin

Fricke J 1992 *Proc. 3rd Int. Symp. Aerogels. J. Non-Cryst. Solids* **145**

Fricke J, Emmerling A 1992 Aerogels—preparation, properties, applications. *Struct. Bonding* **77**, 37–87

Gesser H D, Goswami P C 1989 Aerogels and related porous materials. *Chem. Rev.* **89**, 765–88

Heinrich T, Klett U, Fricke J 1995 Aerogels—nanoporous materials part I: sol–gel process and drying of gels. *J. Porous Mater.* **1**, 7–17

Hüsing N, Schubert U 1998 Aerogels—airy materials: chemistry, structure, and properties. *Angew. Chem. Int. Ed. Engl.* **37**, 23–45

Pajonk G M 1991 Aerogel catalysts. *Appl. Catal.* **72**, 217–66

Pajonk G M 1997 Catalytic aerogels. *Catal. Today* **35**, 319–37

Pekala R W, Hrubesch L W 1995 *Proc. 4th Int. Symp. Aerogels J. Non-Cryst. Solids* **186**

Phalippou J, Vacher R 1998 *Proc. 5th Int. Symp. Aerogels. J. Non-Cryst. Solids* **225**

Schaefer D W 1989 Polymers, fractals, and ceramic materials. *Science* **243**, 1023–7

Schneider M, Baiker A 1995 Aerogels in catalysis. *Catal. Rev. Sci. Eng.* **37**, 515–56

Schneider M, Baiker A 1997 Titania-based aerogels. *Catal. Today* **35**, 339–65

Vacher R, Phalippou J, Pelous J Woignier T 1989 *Proc. 2nd Int. Symp. Aerogels. Rev. Phys. Appl.* **C4**(Suppl. 4)

U. Schubert and N. Hüsing
Technische Universität Wien, Vienna, Austria

Inorganic Nanotube Materials

The original article, which was published earlier, was focused mostly on the structural principles underlying the formation of nanotubes and fullerene-like nanostructures from layered (2-D) compounds (IF). This area witnessed a remarkable upshot over the years with numerous articles appearing in scientific literature, reporting new synthetic strategies which yielded these substances. Not less importantly though, new physical measurements have been carried out and the number of potential applications has increased dramatically.

Generally speaking, the concept of nanotubes and fullerene-like nanoparticles being the most stable phase of layered compounds in the nanodomain, which was proposed and first demonstrated for WS_2 (Tenne *et al.* 1992, see Fig. 1) and MoS_2 (Margulis *et al.* 1993) and their respective selenides, is well established by now. It was verified through numerous experiments and quantum chemical calculations. Using density functional theory (DFT) calculations, Rubio *et al.* proved in 1994 that single-wall BN nanotubes are stable compared to their unfolded analogs. More recently, a combined experimental and DFT tight-binding (DFT–TB) study demonstrated that multi-wall WS_2 (MoS_2) nanotubes of (outer) diameter between 10 nm and 20 nm (4–7 layers) are more stable than any conceivable unfolded structure with the same number of atoms (Seifert *et al.* 2002). In the absence of experimental evidence, DFT calculations proved to be very informative in this respect. Thus, single-wall phosphorus nanotubes were shown to be stable compared to their unfolded analogs (Seifert and Hernandez 2000, Cabria and Mintmire 2004), while phosphorous fullerenes were found to be unstable (Seifert *et al.* 2001). Numerous examples for different inorganic nanotubes have been worked out recently, providing strong evidence for the stability of such systems in the nanoregime. Despite the absence of stable layered type phases, it was argued that compounds exhibiting metastable sp^2-like hybridization, like GaN (Lee *et al.* 1999), or AlN (Zhang and Zhang 2003) can form stable nanotubes and even fullerene-like structures (Chang *et al.* 2001).

New synthetic approaches have been devised to promote the growth of nanotubes from various compounds. Perhaps, the most versatile and successful series of strategies were based on ‘‘chimie douce’’ (soft chemistry), and in particular, the hydrothermal synthesis which was extensively used for the synthesis of numerous new nanotube phases. Thus, $Ce(HPO_4)_2 \cdot H_2O$ nanotubes were synthesized by heating a 6 M solution of phosphoric acid at 50–110 °C and

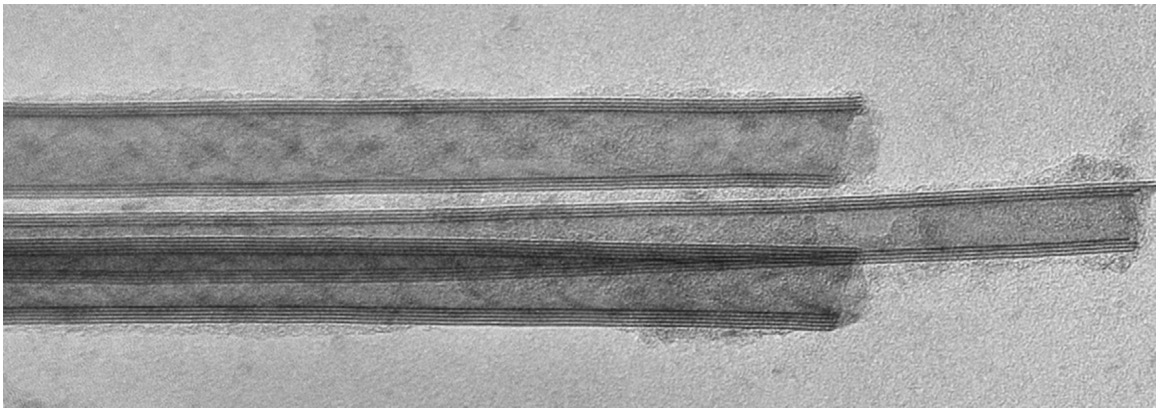


Figure 1
A bundle of three WS_2 nanotubes.

subsequently adding 0.02 M cerium diammonium nitrate solution dropwise with continuous mixing (Tang *et al.* 2004). WO_2Cl_2 nanotubes, the first oxyhalide nanotubes to be synthesized, were obtained by dissolving and stirring the ball-milled WO_2Cl_2 powder in $LiPF_6$ in a nonaqueous solvent for a few days (Armstrong *et al.* 2004). $CuSiO_3 \cdot 2H_2O$ and $Mg_3Si_2O_5(OH)_4$ nanotubes were synthesized by hydrothermal synthesis from an ethanol–water–ammonia solution of the respective nitrate salts and Na_2SiO_3 at 200 °C (Wang *et al.* 2004). Crystals consisting of bundles of single-layer uranyl selenate nanotubules were recently prepared at room temperature by a reaction of uranyl nitrate, butylamine, and H_2SeO_4 in an aqueous solution (Krivovichev *et al.* 2005). Titanate nanotubes have, over the years, received much attention, mainly due to their numerous potential applications. Most notably, Peng and his co-workers found that these nanotubes are formed by first a chemical transformation of the original TiO_2 crystal into the layered $(Na, H)_2Ti_3O_7$ lattice, and then its subsequent exfoliation into molecular sheets of the compound and folding into nanoscrolls (Zhang *et al.* 2003, see also Fig. 2). TiS_2 nanotubes were prepared by the reaction of a nonaqueous solution of Na_2S with $TiCl_4$, followed by solvothermal synthesis at 200 °C (Chen *et al.* 2003a). It should be borne in mind that nanoscrolls, which are halfway in stability between the unstable nanosheets and the fully bonded and stable nanotubes, are a typical product of low-temperature chemistry. This short account presents a few examples from the large body of data dedicated to the low-temperature synthesis of inorganic nanotubes.

Many reports were dedicated to the high-temperature synthesis of inorganic nanotubes, or processes involving a combination of low- and high-temperature synthesis (see, e.g., Li *et al.* 2002). Following the synthesis of nested nanoparticles of WS_2 with fullerene-like structure by the fluidized bed reactor, the same reactor was used for the large-scale synthesis of multiwall WS_2 nanotubes with lengths up to a fraction of a millimeter (Rosentsveig *et al.* 2002). High-temperature synthesis of various metal chalcogenide (MX_2) nanotubes from the MX_3 precursor was reported first by Nath and Rao (2001) and reviewed by Rao and Nath (2003). Chemical vapor transport (CVT) with iodine was used to synthesize single-wall MoS_2 nanotubes with diameter of < 1 nm which self-assemble into hexagonal closed-packing bundles (Remskar *et al.* 2001). The role of iodine in these superstructures has not been fully clarified. CVT has been used also by Chen *et al.* (2003b) to synthesize TiS_2 and other MX_2 nanotubes. WS_2 (Whitby *et al.* 2002) and ReS_2 (Brorson *et al.* 2002) nanotubes were obtained by soaking carbon nanotubes in the respective metal–salt solution and subsequent firing the dried powder at elevated temperatures in H_2S atmosphere. Generally speaking, high-temperature processes yield perfect crystalline nanotubes.

While many nanotubes have been prepared from different kinds of layered compounds, recent efforts were also directed to the synthesis of crystalline nanotubes from quasi-isotropic (3-D) compounds. Such nanotubes are not expected to be fully stable and their surfaces are usually passivated by reaction with the ambient. Most frequently, sacrificial templates are used to endow the lattice with a tubular structure. Thus, GaN nanotubes (zincblende structure with sp^3 bonding) were prepared by chemical vapor deposition of trimethylgallium and ammonia vapor onto ZnO nanowires with hexagonal cross-section at 600–700 °C

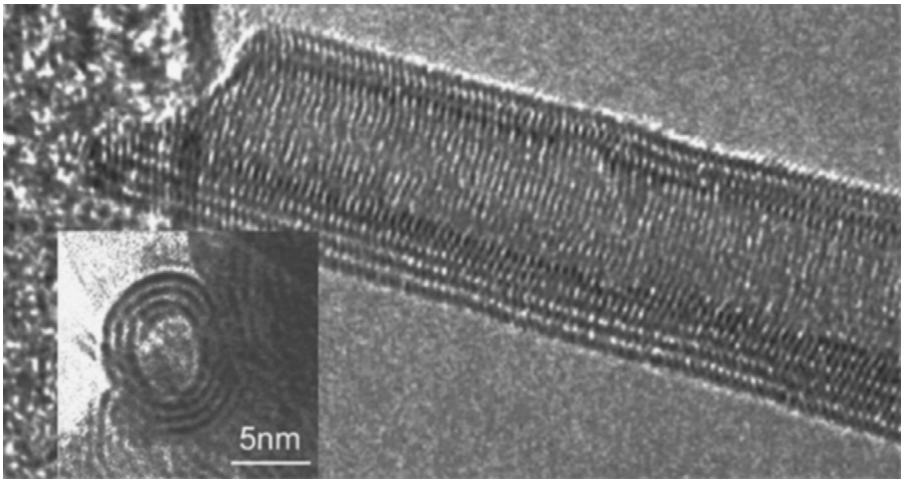


Figure 2
$H_2Ti_3O_7$ nanoscrolls. Courtesy of L. M. Peng.

(Goldeberger *et al.* 2003). The ZnO nanowires were later removed by reductive annealing at 600 °C. An alternative strategy was developed by the Golberg-Bando group (see, e.g., Li *et al.* 2003). Here indium metal and In_2O_3 powders were heated to 1300 °C in the presence of a carbon source (carbothermal synthesis). The product contained crystalline In_2O_3 nanotubes which grew out of the oxide grains. Much of the nanotubes' core was filled with indium. This growth mode suggests that metallic indium droplets served as templates for the vaporized oxide nucleation and growth. This strategy, where the molten droplet was obtained from the oxide by solvothermal reduction using reactive carbon powder[2], was further developed and documented by this group in a number of recent publications. In another study, magnetite (Fe_3O_4) was deposited onto MgO nanowires by pulsed laser ablation process (Liu *et al.* 2004). Subsequently, the MgO template was removed by wet chemical etching in a solution of ammonium sulfate, leaving behind the hollow magnetite nanotubes.

The synthesis of nanotubes from various inorganic compounds permits a systematic study of their properties. The Young's modulus of BN nanotubes was measured to be about 1 TPa (Chopra and Zettl 1998). The mechanical properties of WS_2 nanotubes have been investigated in detail through both, experiment and theory (Kaplan-Ashiri *et al.* 2004, 2005). Tensile and buckling experiments of individual nanotubes were carried out using both atomic force microscope (AFM) and high-resolution scanning electron microscope (HRSEM). The Young's modulus, yield strength, and elongation to yield are 150–170 GPa, 15.7, and $>10\%$, respectively, putting these nanotubes in a unique position of being both very strong yet flexible material. Furthermore, in contrast to macroscopic objects, the mechanical behavior of the nanotubes was shown to be dictated mostly through their intrinsic properties, that is, the strength of the chemical bonds. DFT–molecular dynamic calculations of MoS_2 nanotubes showed that the nanotube is deformed elastically through most of the elongation and its failure is effected through rapid propagation of an atomic defect along the circumference of the nanotube. In a different kind of experiment, bundles of single-wall (3,3) MoS_2 nanotubes were suspended over porous alumina substrate. Using the AFM tip, the nanotubes were deflected vertically into the pores. Using a well-established model and the experimental parameters, the Young's and shear moduli of the nanotube were calculated as being 120 GPa and 0.16 GPa, respectively (Kis *et al.* 2003). The low value of the estimated shear modulus of the MoS_2 nanotube as compared to that of graphite (4.5 GPa) is indicative of its suitability for solid lubrication (*vide infra*).

The effect of shockwave impact on the structural integrity of WS_2 nanotubes (Zhu *et al.* 2002) and MoS_2 nanoparticles with fullerene-like structure (Zhu *et al.* 2005) was evaluated. While the WS_2 nanotubes were stable to impacts of 21 GPa, the MoS_2 were found to withstand impacts as high as 30 GPa with minimal destruction. Furthermore, fullerene-like WS_2 and MoS_2 nanoparticles were recently found to be stable under compression in a diamond anvil cell to pressures of 20 GPa and 30 GPa, respectively (Joly-Pottuz *et al.* 2005). The extraordinary resilience of the nanoparticles against shock and compression offers numerous applications in high-strength materials.

The electronic structure of BN and BC_2N nanotubes was investigated first by the Louie–Cohen group (Rubio *et al.* 1994, Miyamoto *et al.* 1994). BN nanotubes were shown to be large-gap semiconductors with the gap decreasing with the shrinking diameter of the nanotube. BC_2N nanotubes arranged in two different configurations were studied: one a large-gap semiconductor, while in the other, the carbon atoms formed a continuous coil within the hexagonal BN network, giving the nanotubes a metallic character. Seifert *et al.* (2000) carried out a systematic theoretical study of the structure and properties of metal–dichalcogenide nanotubes using DFT-TB calculations. In their earlier papers, they made three important observations: (i) Nanotubes which are derived from semiconducting compounds remain so after being folded into nanotubes. (ii) In contrast to usual nanoparticles, the bandgap of these nanotubes decreases with shrinking diameter of the nanotube. (iii) While (n,n) armchair nanotubes, like their bulk predecessor, exhibit indirect lowest gap, $(n,0)$ zigzag nanotubes exhibit a direct gap. These calculations are so far supported only partially by experimental measurements (see, e.g., Scheffer *et al.* 2002). More recently, the stability and band structure of nanotubes of III–V compounds, like AlN, which generally crystallize in the stable zincblende (diamond-like) structure, was studied. In the present calculations however, a network of sp^2 (graphite-like) bonded aluminum and nitrogen atoms was assumed (see, e.g., Zhao *et al.* 2003). Interestingly, not only were the nanotubes found to be stable against unfolding (though not globally stable), but their electronic structure too was not very different from those of the other inorganic nanotubes as summarized in the three points above.

Various applications have been proposed for the inorganic nanotubes and fullerene-like nanoparticles. The preliminary observation that IF-WS_2 (IF-MoS_2) can serve as high-performance solid lubricants (Rapoport *et al.* 1997) was followed by a detailed investigation that elucidated the complex mechanism underlying the action of these nanoparticles as solid lubricants (for a short summary, see Rapoport *et al.* (2005)). It is believed that under relatively low loads (<0.5 GPa), the main benefit of these nanoparticles emerges from sliding and rolling. Under higher loads, the spacing between the reciprocating pair of surfaces becomes smaller than the size of the nanoparticles, leading thus to their gradual collapse and the transfer of WS_2 nanosheets onto the asperities. Consequently, the WS_2-coated surfaces slide past each other with minimum friction and wear. These findings led to substantial interest in these materials in automobile, aerospace, machining, microelectronics, medical, and other industries. Hence, a major effort for the large-scale commercialization of this technology is now well under way.

The use of inorganic nanotubes for cathode materials in rechargeable batteries has been contemplated. It was shown (Nordlinder *et al.* 2003, Tao *et al.* 2004) that large amounts of the guest atoms can be intercalated within the van der Waals gap of the host nanotubular lattice, giving rise to substantial electrical capacity. Furthermore, the kinetic stability of the nanotubular structures allows for many charge/discharge cycles with only moderate losses in the electrical capacity. Potential applications of inorganic nanotubes in catalysis (Chen *et al.* 2002), hydrogen storage (Chen *et al.* 2003), solar cells (Adachi *et al.* 2003), and in organic light emitting diodes (Qian *et al.* 2004) have been demonstrated. The wide variability of these materials and their rich chemistry offers countless opportunities for a vast number of studies and applications in various technologies.

Bibliography

Adachi M, Murata Y, Okada I, Yoshikawa S 2003 Formation of titania nanotubes and applications for dye-sensitized solar cells. *J. Electrochem. Soc.* **150**, G488–93

Armstrong R A, Canales J, Bruce P G 2004 WO_2Cl_2 nanotubes and nanowires. *Angew. Chem. Intl. Ed.* **43**, 4899–902

Brorson M, Hansen T W, Jacobsen C J H 2002 Rhenium (iv) sulfide nanotubes. *J. Am. Chem. Soc.* **124**, 11582–3

Cabria I, Mintmire J W 2004 Stability and electronic structure of phosphorus nanotubes. *Europhys. Lett.* **65**, 82–8

Chang Ch, Patzer A B C, Sedlmayr E, Steinke T, Sülzle D 2001 Computational evidence for stable inorganic fullerene-like structures of ceramic and semiconductor materials. *Chem. Phys. Lett.* **350**, 399–404

Chen J, Li S, Xu Q, Tanaka K 2002 Synthesis of open-ended MoS_2 nanotubes and the application as the catalyst of methanation. *Chem. Commun.* 1722–3.

Chen J, Tao Z L, Li S L 2003a Lithium intercalation in open-ended TiS_2 nanotubes. *Angew. Chem. Int. Ed. Engl.* **42**, 2147–52

Chen J, Li S L, Tao Z L, Shen Y T, Cui C X 2003b Titanium disulfide nanotubes as hydrogen-storage materials. *J. Am. Chem. Soc.* **125**, 5284–5

Chopra N G, Zettl A 1998 Measurement of the elastic modulus of a multi-wall boron nitride nanotube. *Solid State Commun.* **105**, 297–300

Goldeberger J, He R, Zhang Y, Lee S, Yan H, Choi H J, Yang P 2003 Single crystal GaN nanotubes. *Nature* **422**, 599–602

Joly-Pottuz L, Martin J M, Fleischer N, Tenne R (to be published).

Kaplan-Ashiri I, Cohen S R, Gartsman K, Rosentsveig R, Seifert G, Tenne R 2004 Mechanical behavior of WS_2 nanotubes. *J. Mater. Res.* **19**, 454–9

Kaplan-Ashiri I, Cohen S R, Gartsman K, Ivanovskaya V, Seifert G, Wagner H D, Tenne R 2005 What makes the mechanical properties of (WS_2) nanotubes distinguishable from those of classical macroscopic objects. *Proc. Natl. Acad. Sci.* (in press).

Kis A, *et al.* 2003 Shear and Young's moduli of MoS_2 nanotube ropes. *Adv. Mater.* **15**, 733–6

Krivovichev S V, Kahlenberg V, Tananaev I G, Kaindl R, Mersdorf E, Myasoedov B F 2005 Highly porous uranyl selenate nanotubules. *J. Am. Chem. Soc.* **127**, 1072–3

Lee S M, Lee Y H, Hwang Y G, Elsner, Porezag D, Frauenheim T 1999 Stability and electronic structure of GaN nanotubes from density functional calculations. *Phys. Rev. B* **60**, 7788–91

Li Y, Bando Y, Golberg D 2003 Single crystalline In_2O_3 nanotubes. *Adv. Mater.* **15**, 581–5

Li Y D, Li X L, He R R, Zhu J, Deng Z X 2002 Artificial lamellar mesostructures to WS_2 nanotubes. *J. Am. Chem. Soc.* **124**, 1411–6

Liu Z, Zhang D, Han S, Li C, Lei B, Lu W, Fang J, Zhou C 2004 Single crystalline magnetite nanotubes. *J. Am. Chem. Soc.* **127**, 6–7

Margulis L, Salitra G, Tenne R, Talianker M 1993 Nested fullerene-like structures. *Nature* **365**, 113–4

Miyamoto Y, Rubio A, Cohen M L, Louie S G 1994 Chiral tubules of hexagonal BC_2N. *Phys. Rev. B* **50**, 4976–9

Nath M, Rao C N R 2001 New metal disulfide nanotubes. *J. Am. Chem. Soc.* **123**, 4841–2

Nordlinder S, Lindgren J, Gustafsson T, Edström K 2003 The structure and electrochemical performance of Na^+, K^+, and Ca^{2+} vanadium oxide nanotubes. *J. Electrochem. Soc.* **150**, E280–4

Qian L, Teng F, Jin Z, Zhang Z, Zhang T, Hou Y, Yang S, Xu X 2004 Improved optoelectronic characteristics of light-emitting diodes by using a dehydrated nanotube titanic acid (DNTA)-polymer nanocomposite. *J. Phys. Chem. B* **108**, 13928–31

Rao C N R, Nath M, 2003 Inorganic nanotubes. *Dalton Trans.* 1–25

Rapoport L, Bilik Yu, Feldman Y, Homyonfer M, Cohen S R, Tenne R 1997 Hollow nanoparticles of WS_2 as potential solid-state lubricants. *Nature* **387**, 791–3

Rapoport L, Fleischer N, Tenne R 2005 Applications of WS_2 (MoS_2) inorganic nanotubes and fullerene-like nanoparticles for solid lubrication and for structural nanocomposites. *J. Mater. Chem.* **15**, 1782–8

Remskar M, Mrzel A, Skraba Z, Jesih A, Ceh M, Demsÿar J, Stadelmann P, Levy F, Mihailovic D 2001 Self-assembly of subnanometer-diameter single-wall MoS_2 nanotubes. *Science* **292**, 479–81

Rosentsveig R, Margolin A, Feldman Y, Popovitz-Biro R, Tenne R 2002 WS_2 nanotube bundles and foils. *Chem. Mater.* **14**, 471–3

Rubio A, Corkill J L, Cohen M L 1994 Theory of graphitic boron nitride nanotubes. *Phys. Rev. B* **49**, 5081–4

Scheffer L, Rosentsveig R, Margolin A, Popovitz-Biro R, Seifert G, Cohen S R, Tenne R 2002 Scanning tunneling microscopy study of WS_2 nanotubes. *Phys. Chem. Chem. Phys.* **4**, 2095–8

Seifert G, Hernandez E 2000 Theoretical prediction of phosphorus nanotubes. *Chem. Phys. Lett.* **318**, 355–60

Seifert G, Terrones H, Terrones M, Jungnickel G, Frauenheim T 2000 On the electronic structure of WS_2 nanotubes. *Solid State Commun.* **114**, 245–8

Seifert G, Heine T, Fowler P W 2001 Inorganic nanotubes and fullerenes: structure and properties of hypothetical phosphorus fullerenes. *Eur. Phys. J. D* **16**, 341–3

Seifert G, Ko1hler T, Tenne R 2002 Stability of metal chalcogenide nanotubes. *J. Phys. Chem. B* **106**, 2479–501

Tang C, Bando Y, Golberg D, Ma R 2004 Cerium phosphate nanotubes: synthesis, valence state, and optical properties. *Angew. Chem. Int. Ed. Engl.* **43**, 2–6

Tao Z, Xu L, Gou X, Chen J, Yuana H 2004 TiS_2 Nanotubes as the cathode materials of Mg-ion batteries. *Chem. Commun.* 2081–2

Tenne R, Margulis L, Genut M, Hodes G 1992 Polyhedral and cylindrical structures of WS_2. *Nature* **360**, 444–5

Wang X, Zhang J, Chen J, Kebin Zhou K, Li Y D 2004 Thermally stable silicate nanotubes. *Angew. Chem. Int. Ed. Engl.* **43**, 2017–20

Whitby R L D, Hsu W K, Fearon P K, Billingham N C, Maurin I, Kroto H W, Walton D R M, Boothroyd C B, Firth S, Clark R J H, Collison D 2002 Multiwalled carbon nanotubes coated with tungsten disulfide. *Chem. Mater.* **14**, 2209–17

Zhang D, Zhang R Q 2003 Theoretical prediction on aluminum nitride nanotubes. *Chem. Phys. Lett.* **371**, 426–32

Zhang S, Peng L-M, Chen Q, Du G H, Dawson G, Zhou W Z 2003 Formation mechanism of $H_2Ti_3O_7$ nanotubes. *Phys. Rev. Lett.* **91**, 256103

Zhao M, Xia Y, Zhang D, Mei1 L 2003 Stability and electronic structure of AIN nanotubes. *Phys. Rev. B* **68**, 235–415

Zhu Y Q, *et al.* 2002 Shock-wave resistance of WS_2 nanotubes. *J. Am. Chem. Soc.* **125**, 1329–33

Zhu Y Q, *et al.* 2005 WS_2 and MoS_2 inorganic fullerenes—super shcok absorbers at very high pressures. *Adv. Matter.* **17**, 1500–3

R. Tenne
Department of Materials & Interfaces,
Weizmann Institute, Rehovoth 76100, Israel

L

Laminates: Physical and Mechanical Behavior

A composite laminate is a perfectly bonded assembly of individual plies of a unidirectional composite for which the fibers are parallel in a direction that defines the orientation of the ply within the laminate (Fig. 1 shows a cross-ply laminate having separated the individual plies). Various methods have been used to develop an understanding of damage growth in such laminates. Parvizi *et al.* (1978) developed the first so-called shear lag model of ply cracking in a cross-ply laminate, while Nuismer and Tan (1988) adopted the optimum shear lag approach that is recommended for use. More accurate analyses based on the application of variational techniques to ply crack problems have been given by Hashin (1985, 1987, 1988) and by Nairn (1989, 1995, 1997) and Nairn and Hu (1992). The most accurate methods consider multilayer models of laminates where through-thickness stress and deformation variations are modeled by subdividing each ply of the laminate into subplies. Examples of this approach are the variational approach applied to cross-ply laminates by Pagano (1991), Pagano and Schoeppner (1997), Schoeppner and Pagano (1998), and the technique developed by McCartney (1998b, 2000) for both cross-ply and general symmetric laminates which is consistent with variational mechanics but provides the complete stress and displacement fields within the laminate.

There are three types of behavior that need to be considered when designing composite components. First of all, for relatively low applied loadings, the composite deforms as a linear elastic solid without the formation of any type of microstructural damage. Many models predicting the thermoelastic constants of laminated composites have been developed for this type of behavior (see *Elastic Behavior of Composites**). For larger applied loads, the composite experiences the second type of behavior where microstructural damage formation occurs leading to the progressive degradation of material properties, and to non-linear stress–strain behavior. Engineering components usually exhibit stress concentrations arising from geometrical features such as holes,

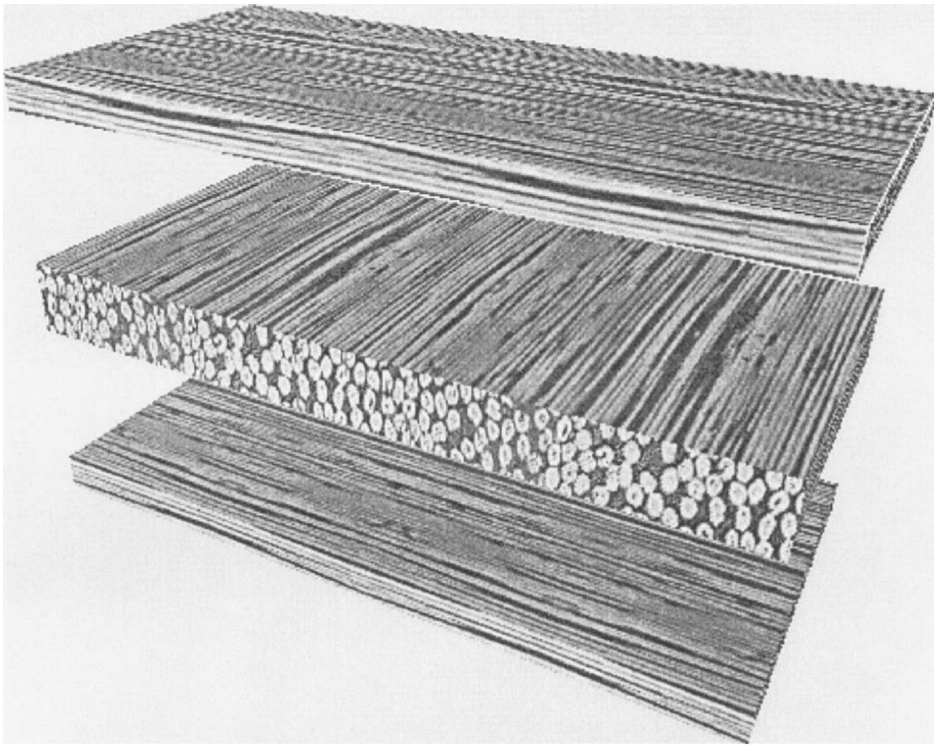


Figure 1
Computer image of the structure of a cross-ply laminate.

---

* Cross references marked by an asterisk are included in this volume.
# Cross references marked by a hash can be found by consulting the Encyclopedia of Materials: Science and Technology.

changes of section, etc. Such concentrations cause localized microstructural damage formation and localized property degradation that leads to the macroscopic redistribution of stress within the component. This is a very important phenomenon that is largely responsible for composite components out-performing their expectations, based on designs using laboratory coupon data for relatively simple loading modes, e.g., uniaxial loading. Following the progressive stable growth of microstructural damage, the third type of behavior ensues where instability and catastrophic failure of the structure occur. Important design objectives are to delay the occurrence of microstructural damage, to extend the progressive microstructural damage growth phase of deformation, and to delay the onset of catastrophic failure thus leading to engineering components that fail "gracefully," i.e., with plenty of warning.

The methodology to be described here is aimed at fulfilling these objectives. It can be applied to any fiber-reinforced composite whose constituents can be regarded as deforming in a linear elastic manner, such as, thermosetting polymer composites and ceramic matrix composites. The methodology can be applied to monotonically loaded, fiber-reinforced thermoplastic and metal composites only if matrix yield and creep are not significant; otherwise, extensions of the methodology must be sought. It can also be applied to fatigue-damaged fiber-reinforced metal laminates when assessing the effect on effective properties of ply cracks that have formed as a result of fatigue loading, and the growth of such damage during cyclic loading.

1. Assumptions of the Methodology

The assumptions on which the methodology is based are:

(i) laminates are symmetric (i.e., ply orientations are balanced about mid-plane of laminate) and constructed of any number of plies that can be made of different anisotropic materials having directional properties, and are perfectly bonded together at all points on the interlaminar interfaces.

(ii) laminates are subject only to loading in the plane of the laminate. Bend deformation of the laminate is not considered.

(iii) there exists a stress-free temperature T_o at which laminates, when in an unloaded state, are free of any internal stresses arising from thermal expansion mismatch between plies. The temperature difference, $\Delta T = T - T_o$, where T is the current temperature, appears in the stress/strain relations and, when non-zero, leads to thermal residual stresses in the composite.

(iv) the undamaged plies of laminates have directional properties and linear stress/strain behavior.

(v) cracks in any ply of the laminate are planar such that the entire ply cross-section is cracked.

2. Undamaged General Symmetric Laminates

In order to characterize the properties of laminates it is useful to introduce orthogonal directions for a ply of the laminate. These are the axial y-direction, the transverse z-direction which is normal to the axial direction lying in the plane of the laminate, and the through-thickness x-direction. The stress and strain components referred to these directions are denoted by σ_{xx}, ε_{xx}, σ_{xy}, ε_{xy}, etc.

2.1 Stress/Strain Relations for Individual Plies in a Laminate

The stress/strain relations for any ply in the *undamaged* laminate whose fibers are parallel to the axial y-direction are of the following form where the superscript, i, is a label denoting the ply being considered

$$\varepsilon^{i}_{xx} = \frac{\sigma^{i}_{xx}}{E^{i}_{t}} - \frac{\nu^{i}_{a}}{E^{i}_{A}}\sigma^{i}_{yy} - \frac{\nu^{i}_{t}}{E^{i}_{T}}\sigma^{i}_{zz} + \alpha^{i}_{t}\Delta T, \tag{1}$$

$$\varepsilon^{i}_{yy} = -\frac{\nu^{i}_{a}}{E^{i}_{A}}\sigma^{i}_{xx} + \frac{\sigma^{i}_{yy}}{E^{i}_{A}} - \frac{\nu^{i}_{A}}{E^{i}_{A}}\sigma^{i}_{zz} + \alpha^{i}_{A}\Delta T, \tag{2}$$

$$\varepsilon^{i}_{zz} = -\frac{\nu^{i}_{t}}{E^{i}_{T}}\sigma^{i}_{xx} - \frac{\nu^{i}_{A}}{E^{i}_{A}}\sigma^{i}_{yy} + \frac{\sigma^{i}_{zz}}{E^{i}_{T}} + \alpha^{i}_{T}\Delta T, \tag{3}$$

$$\varepsilon^{i}_{xy} = \frac{\sigma^{i}_{xy}}{2\mu^{i}_{a}}, \varepsilon^{i}_{xz} = \frac{\sigma^{i}_{xz}}{2\mu^{i}_{t}}, \varepsilon^{i}_{yz} = \frac{\sigma^{i}_{yz}}{2\mu^{i}_{A}} \tag{4}$$

where E_A, E_T, E_t, ν_a, ν_t, ν_A, μ_a, μ_t, μ_A, α_t, α_A and α_T are material properties known as the Young's moduli (E), Poisson's ratios (ν), shear moduli (μ) and thermal expansion coefficients (α) for the undamaged laminate. The upper case suffices A (axial), T (transverse) indicate that the property to which it is attached is for a direction of the laminate lying within the plane of the laminate. Lower case suffixes are attached to properties involving out-of-plane deformation. When the fibers of an individual ply are oriented in the axial direction, and the ply is isotropic in directions normal to the fibers, the stress/strain relations are given elsewhere (see also *Continuous Parallel Fiber Composites: Deformation and Strength**), together with formulae that can be used to estimate the thermoelastic constants of undamaged plies from the fiber and matrix properties.

For other orientations of the plies, where the fibers in the ith ply are parallel and inclined so that the angle between the fiber direction and the y-axis is ϕ_i, the stress/strain relations are of the form (McCartney 2000).

$$\varepsilon^{i}_{xx} = g^{i}_{11}\sigma^{i}_{xx} + g^{i}_{12}\sigma^{i}_{yy} + g^{i}_{13}\sigma^{i}_{zz} + g^{i}_{14}\sigma^{i}_{yz} + \alpha^{i}_{1}\Delta T \tag{5}$$

$$\varepsilon^{i}_{yy} = g^{i}_{12}\sigma^{i}_{xx} + g^{i}_{22}\sigma^{i}_{yy} + g^{i}_{23}\sigma^{i}_{zz} + g^{i}_{24}\sigma^{i}_{yz} + \alpha^{i}_{2}\Delta T \tag{6}$$

$$\varepsilon^{i}_{zz} = g^{i}_{13}\sigma^{i}_{xx} + g^{i}_{23}\sigma^{i}_{yy} + g^{i}_{33}\sigma^{i}_{zz} + g^{i}_{34}\sigma^{i}_{yz} + \alpha^{i}_{3}\Delta T \qquad (7)$$

$$2\varepsilon^{i}_{yz} = g^{i}_{14}\sigma^{i}_{xx} + g^{i}_{24}\sigma^{i}_{yy} + g^{i}_{34}\sigma^{i}_{zz} + g^{i}_{44}\sigma^{i}_{yz} + \alpha^{i}_{4}\Delta T \qquad (8)$$

$$\varepsilon^{i}_{xy} = a^{i}_{11}\sigma^{i}_{xy} + a^{i}_{12}\sigma^{i}_{xz}, \quad \varepsilon^{i}_{xz} = a^{i}_{12}\sigma^{i}_{xy} + a^{i}_{22}\sigma^{i}_{xz} \qquad (9)$$

where, in terms of $m_i = \cos\phi_i$, $n_i = \sin\phi_i$, the coefficients appearing in Eqns. (5)–(9) are:

$$\left.\begin{aligned} &g^{i}_{11} = \frac{1}{E^{i}_{t}}, \quad g^{i}_{12} = -m_i^2\frac{\nu^{i}_{a}}{E^{i}_{A}} - n_i^2\frac{\nu^{i}_{t}}{E^{i}_{T}}, \\ &g^{i}_{13} = -m_i^2\frac{\nu^{i}_{t}}{E^{i}_{T}} - n_i^2\frac{\nu^{i}_{a}}{E^{i}_{A}}, \\ &g^{i}_{14} = 2m_i n_i\left(-\frac{\nu^{i}_{t}}{E^{i}_{T}} + \frac{\nu^{i}_{a}}{E^{i}_{A}}\right) \end{aligned}\right\} \qquad (10)$$

$$\left.\begin{aligned} g^{i}_{22} &= m_i^2 n_i^2\left(\frac{1}{\mu^{i}_{A}} - \frac{2\nu^{i}_{A}}{E^{i}_{A}}\right) + \frac{m_i^4}{E^{i}_{A}} + \frac{n_i^4}{E^{i}_{T}}, \\ g^{i}_{23} &= m_i^2 n_i^2\left(\frac{1}{E^{i}_{A}} + \frac{1}{E^{i}_{T}} - \frac{1}{\mu^{i}_{A}}\right) \\ &\quad -m_i^4\frac{\nu^{i}_{A}}{E^{i}_{A}} - n_i^4\frac{\nu^{i}_{A}}{E^{i}_{A}} \end{aligned}\right\} \qquad (11)$$

$$\left.\begin{aligned} g^{i}_{24} &= m_i n_i(m_i^2 - n_i^2)\left(\frac{1}{\mu^{i}_{A}} - \frac{2\nu^{i}_{A}}{E^{i}_{A}}\right) \\ &\quad -\frac{2m_i^2}{E^{i}_{A}} + \frac{2n_i^2}{E^{i}_{T}}, \\ g^{i}_{33} &= m_i^2 n_i^2\left(\frac{1}{\mu^{i}_{A}} - \frac{2\nu^{i}_{A}}{E^{i}_{A}}\right) + \frac{m_i^4}{E^{i}_{T}} + \frac{n_i^4}{E^{i}_{A}} \end{aligned}\right\} \qquad (12)$$

$$\left.\begin{aligned} g^{i}_{34} &= m_i n_i\left[(m_i^2 - n_i^2)\left(-\frac{1}{\mu^{i}_{A}} + \frac{2\nu^{i}_{A}}{E^{i}_{A}}\right)\right. \\ &\quad \left. + \frac{2m_i^2}{E^{i}_{T}} - \frac{2n_i^2}{E^{i}_{A}}\right], \\ g^{i}_{44} &= \frac{(m_i^2 - n_i^2)^2}{\mu^{i}_{A}} + 4m_i^2 n_i^2\left(\frac{1}{E^{i}_{A}} + \frac{1}{E^{i}_{T}} + \frac{2\nu^{i}_{A}}{E^{i}_{A}}\right) \end{aligned}\right\} \qquad (13)$$

$$\left.\begin{aligned} &\alpha^{i}_{1} = \alpha^{i}_{t}, \quad \alpha^{i}_{2} = m_i^2\alpha^{i}_{A} + n_i^2\alpha^{i}_{T}, \\ &\alpha^{i}_{3} = m_i^2\alpha^{i}_{T} + n_i^2\alpha^{i}_{A}, \quad \alpha^{i}_{4} = 2m_i n_i(\alpha^{i}_{T} - \alpha^{i}_{A}) \end{aligned}\right\} \qquad (14)$$

$$\left.\begin{aligned} a^{i}_{11} &= \frac{1}{2}\left(\frac{m_i^2}{\mu^{i}_{a}} + \frac{n_i^2}{\mu^{i}_{t}}\right), \\ a^{i}_{12} &= \tfrac{1}{2}m_i n_i\left(\frac{1}{\mu^{i}_{t}} - \frac{1}{\mu^{i}_{a}}\right), \\ a^{i}_{22} &= \frac{1}{2}\left(\frac{m_i^2}{\mu^{i}_{t}} + \frac{n_i^2}{\mu^{i}_{a}}\right) \end{aligned}\right\} \qquad (15)$$

2.2 *Macroscopic Effective Stress/Strain Relations for Any Undamaged Symmetric Laminate*

A symmetric laminate has a ply geometry that is symmetric about the mid-plane of the laminate. Consider a laminate of total thickness $2h$ comprising $2N$ plies. The plies in one half of the laminate are labelled $i = 1 \ldots N$. The characteristic property of undamaged symmetric laminates subject to in-plane loading is that the axial, transverse and in-plane shear strains, ε, ε_T and γ respectively, are uniform throughout those parts of the laminate that are not affected by the presence of laminate edges, and that the through-thickness stress, σ_t, is also uniform in all plies of the laminate. In addition the stress components σ_{xy}, σ_{xz} are zero everywhere. When the laminate as a whole is considered the stress/strain relations (Eqns. (5)–(8)) are applied to each ply in the laminate such that $\varepsilon^{i}_{yy} = \varepsilon$, $\varepsilon^{i}_{zz} = \varepsilon_T$, $2\varepsilon^{i}_{yz} = \gamma$, $\sigma_{xx} = \sigma_t$. The effective applied axial, transverse and in-plane shear stresses σ^{i}_{yy}, σ^{i}_{zz}, τ^{i}_{yz} will be uniform in each ply but will be discontinuous at ply interfaces.

The effective axial stress σ, transverse stress σ_T, and in-plane shear stress τ are defined by total normal and shear loads acting on the laminate edges divided by the total area of cross-section of the edges over which the load is acting. The through-thickness displacement on the laminate faces is divided by the half-thickness of the laminate when defining the through-thickness strain ε_t. It is then possible to calculate the effective stresses σ, σ_T and τ in each ply together with the effective through-thickness strain ε_t using:

$$\begin{aligned} &\sigma = \frac{1}{h}\sum_{i=1}^{N} h_i\sigma^{i}_{yy}, \quad \sigma_T = \frac{1}{h}\sum_{i=1}^{N} h_i\sigma^{i}_{zz}, \\ &\tau = \frac{1}{h}\sum_{i=1}^{N} h_i\sigma^{i}_{yz}, \quad \varepsilon_t = \frac{1}{h}\sum_{i=1}^{N} h_i\varepsilon^{i}_{t} \end{aligned} \qquad (16)$$

where ε^{i}_{t} is the uniform through-thickness strain in the i^{th} ply. These values are then used to calculate the effective stresses σ, σ_T and τ and through-thickness strain ε_t for the laminate as a whole. Stress/strain relations for the laminate may then be derived (McCartney 2000) which are of the same form as the relations (Eqns. (5)–(8)) assumed for individual ply properties.

The laminate stress/strain relations are written in the following form that generalizes the notation given in Eqns. (1)–(4):

$$\varepsilon_t = \frac{\sigma_t}{E_t} - \frac{\nu_a}{E_A}\sigma - \frac{\nu_t}{E_T}\sigma_T - \frac{\lambda_t}{E_A}\tau + \alpha_t\Delta T \qquad (17)$$

$$\varepsilon = -\frac{\nu_a}{E_A}\sigma_t + \frac{\sigma}{E_A} - \frac{\nu_a}{E_A}\sigma_T - \frac{\lambda_A}{E_A}\tau + \alpha_A\Delta T \qquad (18)$$

$$\varepsilon_T = -\frac{\nu_t}{E_T}\sigma_t - \frac{\nu_A}{E_A}\sigma + \frac{\sigma_T}{E_T} - \frac{\lambda_T}{E_A}\tau + \alpha_T\Delta T \qquad (19)$$

$$\gamma = -\frac{\lambda_t}{E_A}\sigma_t - \frac{\lambda_A}{E_A}\sigma - \frac{\lambda_T}{E_A}\sigma_T + \frac{\tau}{\mu_A} + \alpha_S\Delta T \qquad (20)$$

The parameters λ_A, λ_T and λ_t are ratios associated with the effect of shear stress on axial and transverse strain respectively. The parameter α_S is the thermal expansion coefficient associated with the effect of temperature changes on the shear strain.

3. *Effects of Ply Cracking on Thermoelastic Constants of Damaged Laminates*

The first damage mode encountered when laminates are subject to multi-axial loading is the formation of ply cracks in plies whose fiber direction is transverse to the principal loading direction. Ply cracks usually form suddenly and completely crack the ply cross-section. Increasing loads lead to the progressive formation of ply crack arrays. When a ply crack forms in an undamaged laminate the load it carried is transferred to neighboring plies which are intact as the fibers are usually in a different direction to those in the cracked ply. The stress transfer occurs through the action of shear tractions on the interfaces between the cracked ply and its neighbors. The nonuniform stress field that results has a significant effect on the values of the thermoelastic constants of the laminate.

Recently, ply crack formation has been studied in quasi-isotropic laminates, (Crocker *et al.* 1997, McCartney 1998b, 1999, 2000) including finite element analysis of cracked laminates for the case of generalized plane strain deformation (Tong *et al.* 1997). For laminated systems, a methodology has been developed (McCartney 1999, 2000) that is capable of predicting accurately the stress and displacement distributions at all points in a cross-ply laminate containing arrays of ply cracks. Such solutions have been used to predict the dependence of the laminate properties, such as Young's moduli, Poisson's ratios, and thermal expansion coefficients, on the degree of cracking, and more importantly, the dependence of these properties on the state of loading. The important result of this analysis is that the stress/strain relations for a damaged laminate (where all ply cracks are stress-free) are of the form given by Eqns. (17)–(20). The occurrence of ply cracking does not alter the form of the stress/strain relations. Only the values of the thermoelastic constants are affected by the damage.

3.1 *Macroscopic Effective Stress–Strain Relations for a Laminate*

Consider a general symmetric laminate subject to multi-axial loading whose damage, in the form of stress-free ply cracks (as illustrated in Fig. 2 for the case of a cross-ply laminate), is characterized by the damage parameter ω. This parameter is a single label that is used to represent a series of damage modes (and associated variables), for example the effective crack densities in each ply. It is assumed that the distribution of damage is homogeneous when viewed at the macroscopic level. When any physical parameter f is damaged its value is denoted by $f(\omega)$ and when undamaged it is denoted by f. In order to simplify the analysis it is assumed that ply cracking occurs only in the 90° plies of a general symmetric laminate for which any ply (in one half of the laminate) having orientation θ has a corresponding -θ ply in the same half of the ply made of the same material and having the same thickness as the θ ply. For this special case the shear coupling terms λ_t , λ_A and λ_T, and the expansion coefficient α_S, in the stress–strain equations for a damaged laminate may be neglected. As a result, the shear stress–strain relations for a damaged and undamaged laminate have respectively the following simple forms:

$$\tau = \mu_A(\omega)\gamma, \quad \tau = \mu_A \gamma \tag{21}$$

McCartney (2000) has considered more general laminates where shear coupling terms are included in the analysis. For the special case when in-plane shear deformation is absent, the effective stress/strain relations of the *damaged* laminate may be expressed in the form:

$$\varepsilon_t = \frac{\sigma_t}{E_t(\omega)} - \frac{\nu_a(\omega)}{E_A(\omega)}\sigma - \frac{\nu_t(\omega)}{E_T(\omega)}\sigma_T + \alpha_t(\omega)\Delta T \tag{22}$$

$$\varepsilon = -\frac{\nu_a(\omega)}{E_A(\omega)}\sigma_t + \frac{\sigma}{E_A(\omega)} - \frac{\nu_A(\omega)}{E_A(\omega)}\sigma_T + \alpha_A(\omega)\Delta T \tag{23}$$

$$\varepsilon_T = -\frac{\nu_t(\omega)}{E_T(\omega)}\sigma_t - \frac{\nu_A(\omega)}{E_A(\omega)}\sigma + \frac{\sigma_T}{E_T(\omega)} + \alpha_T(\omega)\Delta T \tag{24}$$

The values of the thermoelastic constants in Eqns. (22)–(24) may be determined by detailed stress analysis (McCartney 2000) for any distribution of ply cracks. It has been shown, by considering the crack closure conditions for a laminate subject to uniaxial loading in the axial, in-plane transverse, and though-thickness directions, that:

$$\frac{\alpha_t(\omega) - \alpha_t}{\dfrac{\nu_a}{E_A} - \dfrac{\nu_a(\omega)}{E_A(\omega)}} = \frac{\alpha_A(\omega) - \alpha_A}{\dfrac{1}{E_A(\omega)} - \dfrac{1}{E_A}} = \frac{\alpha_T(\omega) - \alpha_T}{\dfrac{\nu_A}{E_A} - \dfrac{\nu_A(\omega)}{E_A(\omega)}} = k_1 \tag{25}$$

$$\begin{aligned} \frac{\dfrac{\nu_t}{E_T} - \dfrac{\nu_t(\omega)}{E_T(\omega)}}{\dfrac{\nu_a}{E_A} - \dfrac{\nu_a(\omega)}{E_A(\omega)}} &= \frac{\dfrac{\nu_A}{E_A} - \dfrac{\nu_A(\omega)}{E_A(\omega)}}{\dfrac{1}{E_A(\omega)} - \dfrac{1}{E_A}} = \frac{\dfrac{1}{E_T(\omega)} - \dfrac{1}{E_T}}{\dfrac{\nu_A}{E_A} - \dfrac{\nu_A(\omega)}{E_A(\omega)}} \\ &= \frac{\alpha_T(\omega) - \alpha_T}{\alpha_A(\omega) - \alpha_A} = k \end{aligned} \tag{26}$$

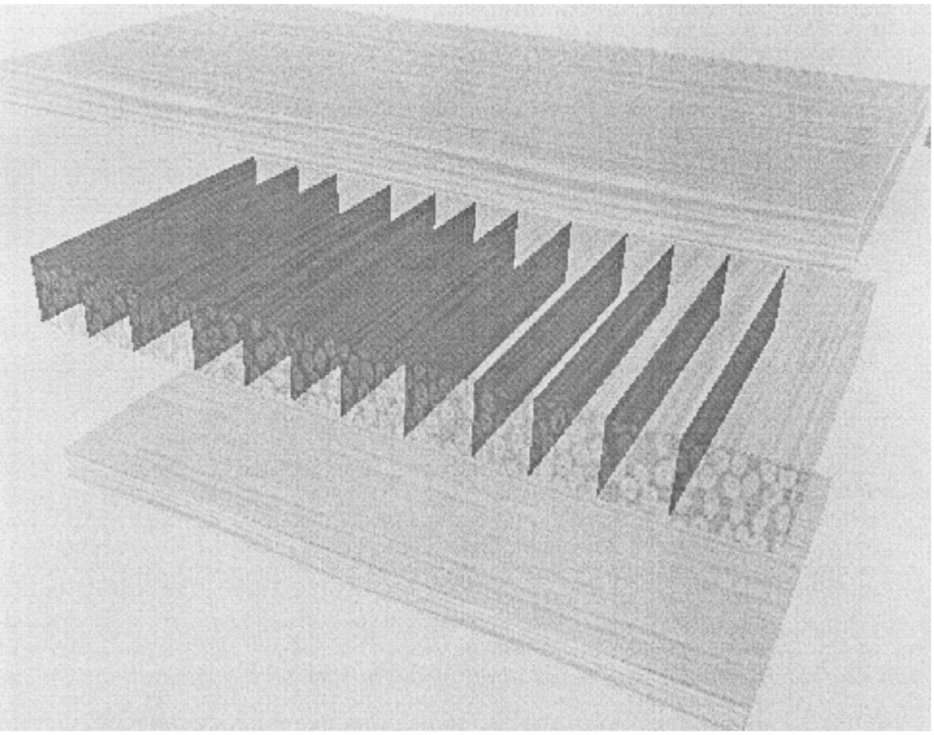


Figure 2
Computer image of the ply cracking that occurs in a cross-ply laminate during loading.

$$\frac{\dfrac{1}{E_t(\omega)}-\dfrac{1}{E_t}}{\dfrac{\nu_a}{E_A}-\dfrac{\nu_a(\omega)}{E_A(\omega)}}=\frac{\dfrac{\nu_a}{E_A}-\dfrac{\nu_a(\omega)}{E_A(\omega)}}{\dfrac{1}{E_A(\omega)}-\dfrac{1}{E_A}}=\frac{\dfrac{\nu_t}{E_T}-\dfrac{\nu_t(\omega)}{E_T(\omega)}}{\dfrac{\nu_A}{E_A}-\dfrac{\nu_A(\omega)}{E_A(\omega)}}=\frac{\alpha_t(\omega)-\alpha_t}{\alpha_A(\omega)-\alpha_A}=k' \tag{27}$$

where

$$k_1=\frac{E_A[\alpha_A+\mathrm{B}\alpha_T-\mathrm{C}]}{1-\nu_A\mathrm{B}},\quad k=\frac{E_A}{E_T}\frac{\mathrm{B}-\nu_A\dfrac{E_T}{E_A}}{1-\nu_A\mathrm{B}},\quad k'=\frac{E_A\mathrm{A}-\nu_a-\nu_t\dfrac{E_A}{E_T}\mathrm{B}}{1-\nu_A\mathrm{B}} \tag{28}$$

and where A, B and C are constants calculated from unidirectional undamaged ply properties (in bold).

$$\mathrm{A}=\frac{\boldsymbol{\nu_t}}{\boldsymbol{E_T}}+\frac{\boldsymbol{\nu_a\nu_A}}{\boldsymbol{E_A}},\ \mathrm{B}=\boldsymbol{\nu_A},\ \mathrm{C}=\boldsymbol{\alpha_T}+\boldsymbol{\nu_A\alpha_A}.$$

The parameters k_1, k and k' are laminate constants that are independent of the state of damage.

3.2 Thermo-elastic Constants for Damaged Laminates

It can be shown from Eqns. (25)–(27) that the thermoelastic constants of a damaged laminate are of the form (McCartney 2000)

$$\left.\begin{aligned}\frac{1}{E_A(\omega)}-\frac{1}{E_A}&=\Phi(\omega),\\ \frac{1}{E_T(\omega)}-\frac{1}{E_T}&=k^2\Phi(\omega),\\ \frac{1}{E_t(\omega)}-\frac{1}{E_t}&=(k')^2\Phi(\omega)\end{aligned}\right\} \tag{29}$$

$$\left.\begin{aligned}\frac{\nu_a}{E_A}-\frac{\nu_a(\omega)}{E_A(\omega)}&=k'\Phi(\omega),\\ \frac{\nu_A}{E_A}-\frac{\nu_A(\omega)}{E_A(\omega)}&=k\Phi(\omega),\\ \frac{\nu_t}{E_T}-\frac{\nu_t(\omega)}{E_T(\omega)}&=kk'\Phi(\omega)\end{aligned}\right\} \tag{30}$$

$$\left.\begin{aligned}\alpha_A(\omega)-\alpha_A&=k_1\Phi(\omega),\\ \alpha_T(\omega)-\alpha_T&=kk_1\Phi(\omega),\\ \alpha_t(\omega)-\alpha_t&=k'k_1\Phi(\omega)\end{aligned}\right\} \tag{31}$$

where all the effective thermoelastic constants of a damaged laminate can be calculated in terms of the single damage dependent parameter $\Phi(\omega)$ that characterises the change in the values of all thermoelastic constants resulting from damage formation. They enable an exact and simple means of fully characterizing the effects of microstructural damage on the effective thermoelastic constants of laminates, requiring only the value of the parameter $\Phi(\omega)$ which depends upon the axial modulus, that is frequently measured in practice or can be estimated from a suitable model. It is emphasizd that the above relations can be derived without having to solve for the stress and displacement distributions in a cracked laminate. Such solutions are needed only to derive values for the effective thermoelastic constants. Nairn (1997) has reached a similar conclusion for a limited set of elastic constants in the case of cross-ply laminates. The results are in fact valid for any damage in laminates where the damage is in the form of cracks or delaminations that are stress-free. The results can also be applied to the degraded properties of laminates that arise during fatigue loading.

4. Predicting Progressive Ply Crack Formation in Multiple-ply Laminates

Consider a laminate that already contains ply cracks in some (but not necessarily all) 90° plies such that the damage is characterized by the parameter ω_1. Then consider the simultaneous formation of new ply cracks in the 90° plies such that the resulting damage state is characterized by the damage parameter ω_2. The locations of the ply cracks are assumed to be such that the overall deformation of the laminate is governed by the effective stress/strain equations (Eqns. (22)–(24)). This means that the crack distribution is regarded as being sufficiently uniformly distributed for shear and out-of-plane bending deformations to be negligible. The new crack surfaces are assumed to form quasi-statically under conditions of fixed effective applied tractions and fixed temperature. From energy balance considerations and the fact that kinetic energy is never negative, the criterion for crack formation under these conditions has the form (McCartney 1998a)

$$\Delta\Gamma + \Delta G < 0 \tag{32}$$

where the energy absorbed in volume V of laminate by the formation of the new cracks is given by

$$\Delta\Gamma = V[\Gamma(\omega_2) - \Gamma(\omega_1)] \tag{33}$$

In Eqn. (33) the parameter $\Gamma(\omega)$ denotes the energy absorbed in unit volume of laminate during the formation of new crack surfaces in the 90° plies that have led to the damage state ω. The energy absorbed is taken as the area of crack surface formed multiplied by the fracture energy for transverse cracking in the 90° plies. The corresponding change of Gibbs free energy (equivalent to complementary energy which must be used when crack growth occurs under fixed applied loads) in the region V of the laminate is

$$\Delta G = \int_{\mathrm{V}} [g(\omega_2) - g(\omega_1)]\mathrm{d}V \tag{34}$$

It has been shown (McCartney 2000) that

$$\frac{1}{V}\int_{V} [g(\omega) - g]\mathrm{d}V = -\tfrac{1}{2}\Phi(\omega)[s - s_{\mathrm{o}}]^2 \tag{35}$$

where s is the effective applied stress for multiaxial loading defined by

$$s = k'\sigma_{\mathrm{t}} + \sigma + k\sigma_{\mathrm{T}} \tag{36}$$

and where $s_{\mathrm{o}} = -k_1\Delta T$ is the ply crack closure stress when the laminate is subject only to uniaxial loading in the axial direction normal to the ply cracks in the 90° plies.

The absorbed energy function $\Gamma(\omega)$, appearing in Eqn. (33), is calculated for length $2L$ of laminate using:

$$\Gamma(\omega) = \frac{h_1}{hL}\sum_{\mathrm{j}=1}^{m} \delta_{\mathrm{j}}\gamma_{\mathrm{j}} \tag{37}$$

where $2h_1$ is the total thickness of all 90° plies in the laminate having total thickness $2h$, and where m is the number of potential cracking sites in 90° plies which are ordered in a regular way, e.g., from top to bottom in the plies which are taken in order from the center of the laminate to the outside, symmetry about the mid-plane of the laminate being assumed. The quantity $2\gamma_{\mathrm{j}}$ is the fracture energy for the growth of a ply crack in the j^{th} potential cracking site of the 90° plies. The parameters δ_{j} describe the crack pattern in the laminate such that:

$$\delta_{\mathrm{j}} = \begin{cases} 0 & \text{if } \mathrm{j}^{\mathrm{th}} \text{ site of the } 90^\circ \text{ ply is uncracked,} \\ 1 & \text{if } \mathrm{j}^{\mathrm{th}} \text{ site of the } 90^\circ \text{ ply is cracked} \end{cases} \tag{38}$$

It is of great practical interest to know the level of applied loading that is likely to cause the occurrence of ply cracking. Ply crack formation can lead to a loss of stiffness, an important issue for stiffness critical components, or to leakage in the case of liquid containers and pressure vessels. It follows from Eqn. (29) and Eqns. (32)–(35) that the formation of an initial set of ply cracks occurs when:

$$s > \sqrt{\left(\frac{2\Gamma(\omega)}{\dfrac{1}{E_{\mathrm{A}}(\omega)} - \dfrac{1}{E_{\mathrm{A}}}}\right)} + s_{\mathrm{o}} \tag{39}$$

The result (Eqn. (39)) is consistent with fracture mechanics where the energy balance associated with a long ply crack propagating across the laminate is considered. It is also consistent with the results of Pagano *et al.* (1998) who applied the three dimensional J-integral to the tip region of a long ply crack in a cross-ply laminate.

It is useful to discuss progressive cracking in a length $2L$ of laminate where the damage parameter ω is defined by $\{L_1, L_2, \ldots L_n\}$ where n is the number of cracks in length $2L$ of the 90° plies and where L_i, $I = 1 \ldots n$ denote the distances between neighboring cracked planes. The criterion for progressive crack formation may be written in the form:

$$s > \sqrt{\left(\frac{2[\Gamma(\omega_2) - \Gamma(\omega_1)]}{\dfrac{1}{E_A(\omega_2)} - \dfrac{1}{E_A(\omega_1)}}\right)} + s_o \qquad (40)$$

It is emphasised that the exceedingly simple results (Eqns. (39) and (40)) are exact, and that they take full account of the effects of anisotropy, of multi-axial loading and of the effect of thermal residual stresses. The results involve only macroscopic properties defined at the laminate level even though they have been derived from a detailed consideration of the stress and deformation distributions in a cracked laminate at the ply level. The results are based on the single assumption of damage homogeneity where damage formation in the laminate is such that the stress/strain relationships (Eqns. (22)–(24)) are always satisfied.

When using Eqn. (40) to investigate progressive crack formation (where cracks form one at a time) the fracture energies $2\gamma_j$ are taken at random from a statistical distribution of fracture energies (assumed to be normal) and then allocated at random to the various potential fracture sites in the 90° plies. The value of the applied stresses are then regarded as slowly increasing until the condition (Eqn. (40)) can be met at just one potential fracture site in the laminate. This means that crack formation must be investigated at all sites in order to determine the location of the first crack to form. This procedure has to be repeated for every new crack that forms during loading. If the fracture energies at each site have the same value, then the criterion (Eqn. (40)) predicts that the crack densities are always uniform having values ρ_o, $2\rho_o$, $4\rho_o$, etc., where ρ_o is the initial crack density. Thus new crack formation for this case occurs as a progressive doubling of the crack density at discrete applied stress levels.

5. *Thick Laminates*

When the thickness of the laminate becomes of the same order as the laminate length and width, the laminate exhibits behavior that cannot adequately be represented by classical laminate theory. The reason for this behavior arises from edge effects in laminates which lead to stress transfer at ply interfaces near the edges of the laminate. Classical laminate theory is based on the assumption that each ply experiences the same in-plane strains and rotations leading to axial and in-plane transverse stress distributions that are discontinuous at ply interfaces. Such stress discontinuities cannot apply at stress-free edges so that stress transfer between plies through the action of interfacial shear tractions must occur in order to satisfy the stress-free condition at laminate edges.

While much literature has been devoted to the development of higher order deformation theories, a good approach is to attempt to model the stress transfer in the way that has already been achieved for stress transfer in cracked laminates (McCartney 2000). Edge effects can be modeled by predicting stress transfer at traction-free laminate edges using these ply crack models.

6. *Strength Issues*

For laminates, the strength is the maximum load that can be supported by the laminate before the occurrence of catastrophic failure where a test specimen separates into two pieces. While the accurate measurement of strength is often difficult, its prediction is even more challenging, as many damage modes occur before the catastrophic failure event, such as ply cracking, delamination and fiber fracture, and the associated fiber/matrix debonding. Experience has shown that the strength of the laminate is usually determined by the strength of the plies whose fibers are aligned in the direction of principal loading. This is particularly true for systems where fiber/matrix debonding and delamination occur. Indeed for ceramic matrix composites, where there are weak fiber/matrix interfaces, matrix cracking can occur to the extent that the matrix is virtually unloaded and the load is taken wholly by the fibers. Such a damaged composite is then behaving as a loose bundle of fibers. For strongly bonded composites this is not the case as fiber fractures lead to the growth of dominant cracks through neighboring matrix and fibers, resulting in much lower strengths.

For laminated polymer composites the situation is more complex. Monte Carlo modeling can be used to estimate the strength of unidirectional composites (see also: *Continuous Parallel Fiber Composites: Deformation and Strength**). Such modeling needs to be introduced into laminate calculations by applying the Monte Carlo model to the plies in which fibers are parallel to the principal loading direction. The effects of ply cracking can either be ignored (upper bound prediction of strength) or taken into account approximately by applying to the Monte Carlo model a larger stress that accounts for the stress transfer from the cracked plies to the plies where fibers are parallel

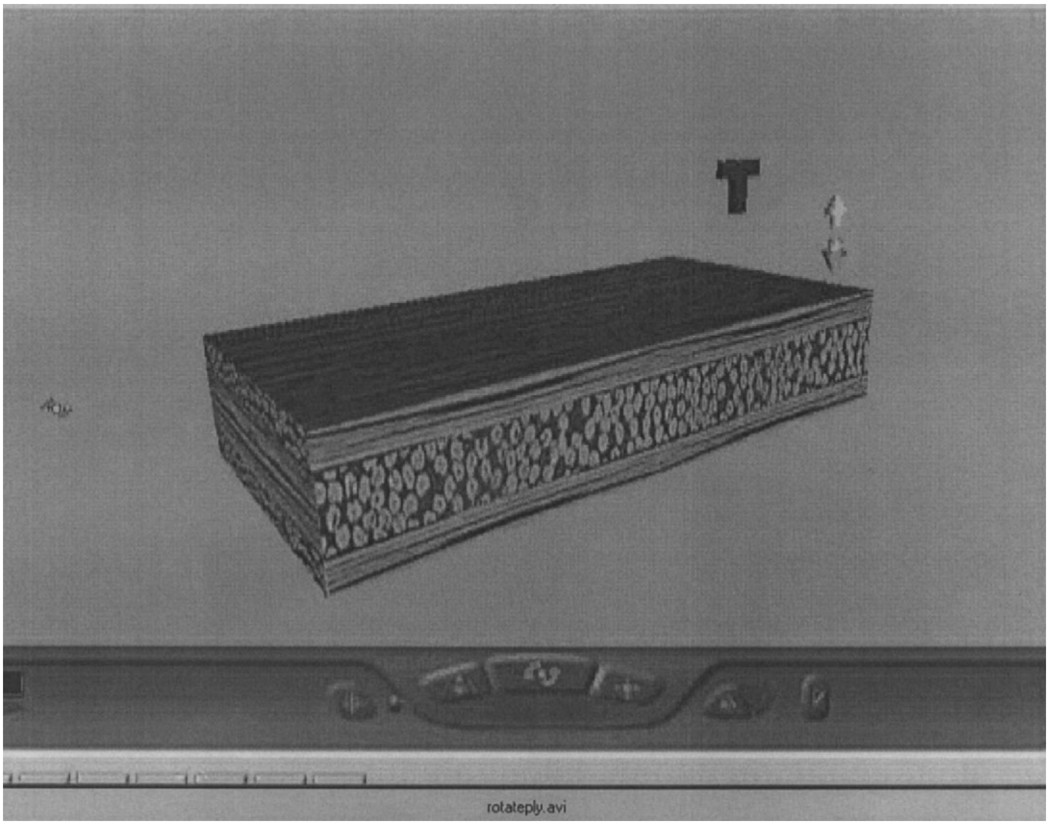


Figure 3
Image from a video showing the operation of the web-based demonstration of ply cracking in a cross-ply laminate (courtesy of N. McCormick, NPL, UK).

to the principal loading direction. Such developments are in progress for the case of cross-ply laminates.

7. Application of Models in Design

The models described in this article may be used to perform the following tasks that can be exploited at some stage during the design of composite structural elements:

Predict undamaged properties of general symmetric laminates from the properties and orientations of the individual plies.

Predict the conditions for the initial formation of fully developed ply cracks in a general symmetric laminate subject to general in-plane loading and thermal residual stresses. In components subject to fatigue loading, designers will want to avoid any form of microstructural damage.

Predict the progressive formation of ply cracks in a single orientation during the monotonic loading of a general symmetric laminate. Multi-axial nonlinear stress–strain behavior is predicted (needed to understand strain softening due to microstructural damage), and the progressive degradation of the thermoelastic constants as a function of applied stress or strain. Such nonlinear behavior leads to load transfer in a structure from regions of stress concentration, where microstructural damage forms and degrades properties, to other regions in the structure. This phenomenon can lead to composite components performing beyond expectations using designs based on coupon data and the neglect of damage effects.

Predict the effects of varying the thickness of the plies in a laminate. Ply cracking is well known to become more difficult when the thickness of the plies is reduced. The energy based approach of this article can model this behavior, contrasting sharply with stress based approaches. Doubling the thickness and applied loads does lead to a different damage evolution that is not predicted by strength based methods.

Predict the effects of temperature changes on ply crack formation (can be used to investigate thermal cracking following manufacture, or during cryogenic cooling).

8. Web Access to Laminate Models

The ply cracking model for cross-ply and general symmetric laminates can be accessed over the internet using the address, http://materials.npl.co.uk. Following registration users of the demonstration models can predict progressive ply crack formation during loading and its effect on properties. Three-dimensional visualization software enables the user to examine the geometry of any laminate, observe damage formation during loading, and plot graphs of results. By double-clicking on Fig. 3 a video is shown that illustrating some of the features that can be operated over the internet.

Bibliography

Crocker L E, Ogin S L, Smith P A, Hill P S 1997 Intra-laminar fracture in angle-ply laminates. *Composites A* **28A**, 839–46

Hashin Z 1985 Analysis of cracked laminates: a variational approach. *Mech. of Mater*. **4**, 121–36

Hashin Z 1987 Analysis of orthogonally cracked laminates under tension. *J. Appl. Mech*. **54**, 872–9

Hashin Z 1988 Thermal expansion coefficients of cracked laminates. *Comp. Sci. Tech*. **31**, 247–60

McCartney L N 1998a Prediction of microcracking in composite materials. In: Cherepanov G P (ed.) *Fracture: A Topical Encyclopaedia of Current Knowledge Dedicated to Alan Arnold Griffith*. Krieger Publishing, Melbourne, USA, pp. 905–16

McCartney L N 1998b Predicting transverse crack formation in cross-ply laminates resulting from microcracking. *Comp. Sci. Tech*. **58**, 1069–81

McCartney L N 1999 An effective stress controlling progressive damage formation in laminates subject to triaxial loading. *Proc. Conf. Deformation and Fracture in Composites*. IoM Communications, London, pp. 23–32

McCartney L N 2000 Model to predict effects of triaxial loading on ply cracking in general symmetric laminates. *Comp. Sci. Tech*. **60**, 2255–79

Nairn J A 1989 The strain energy release rate of composite microcracking: a variational approach. *J. Comp. Mater*. **23**, 1106–29

Nairn J A 1995 Some new variational mechanics results on composite microcracking. *Proc 10th Int. Conf. Composite Materials*. Wordhead Publishing Ltd., Cambridge, UK, Whistler, BC, Canada, pp. 423–30

Nairn J A 1997 Fracture mechanics of composites with residual thermal stresses. *J. Appl. Mech*. **64**, 805

Nairn J A, Hu S 1992 The formation and effect of outer-ply microcracks in cross-ply laminates: a variational approach. *Eng. Fract. Mech*. **41**, 203–21

Nuismer R J, Tan S C 1988 Constitutive relations of a cracked composite lamina. *J. Comp. Mater*. **22**, 306–21

Parvizi A, Garrett K W, Bailey J E 1978 Constrained cracking in glass fiber reinforced epoxy cross-ply laminates. *J. Mater. Sci*. **13**, 195–201

Pagano N J 1991 Axisymmetric micromechanical stress fields in composites. *Proc. IUTAM Symp. Local Mechanics Concepts for Composite Material Systems*. Springer-Verlag, Berlin, pp. 1–26

Pagano N J, Schoeppner G A 1997 Some transverse cracking problems in cross-ply laminates. Proc. AIAA *Conf.* Orlando, FL.

Pagano N J, Schoeppner G A, Kim R, Abrams F 1998 Steady state cracking and edge effects in thermo-mechanical transverse cracking of cross-ply laminates. *Comp. Sci. Tech*. **58**, 1811–25

Schoeppner G A, Pagano N J 1998 Stress fields and energy release rates in cross-ply laminates. *Int. J. Solids Struct*. **35**, 1025–55

Tong J, Guild F J, Ogin S L, Smith P A 1997 On matrix crack growth in quasi-isotropic laminates–I. Experimental investigation–II Finite element analysis. *Comp. Sci. Tech*. **57**, 1527–45

L. N. McCartney

National Physical Laboratory, Teddington, UK

Liquid Crystalline Polymers, Dispersed

Liquid crystal–polymer composites have been studied intensively for both fundamental aspects of science and practical applications. The main interest in the research and industrial development of “plastic” liquid crystal technology originated on the premise of *stabilizing* liquid crystals through their interaction with an underlying polymer matrix and for manufacturing a new generation of liquid crystal display devices. Although the concentration of one of the components in a composite may be small, the resultant effect on the characteristic electro-optic response is usually large. Based on their morphology, there are two main types of liquid crystal–polymer composites. The polymer dispersed liquid crystal (PDLC) represents one end of the spectrum, where a high concentration (greater than 20%) of polymer is used to create a macroscopically phase-separated system of continuous polymer phase with small liquid crystal droplets embedded in it. This results in the technology of large-area self-supporting flexible electro-optic displays. The addition of small amounts of polymer (less than 10%) to a liquid crystal host falls into the regime of polymer stabilized liquid crystal (PSLC). Here the phase-separating polymer forms a network of narrow strands in the continuous liquid crystal matrix, which are aligned along the director and thus anchor this orientation throughout the working cycle of the device. The principal technological advantage here is in reducing the characteristic length and the time scale of the electro-optic response.

The early milestones of the research in confined liquid crystals include studies of director orientation and response to external (electric) field in narrow capillary tubes and spherical droplets. Characteristic topological defects (linear disclinations, point-like monopoles and surface boojums, see Fig. 1) and their transformations have been observed, causing much excitement from theoretical physicists and applied mathematicians. Remarkably universal properties and conservation laws of topological defects have led to significant advances in other areas, reaching as far as the mechanisms of galaxy formation at the early stages of expansion of the Universe (e.g., Zurek 1996) (see *Nematic Liquid Crystals: Defects*[#]). However, up to the mid-1980s, confined liquid crystals were only curious model systems from the basic science perspective.

The practical usefulness of spherical nematic droplets was appreciated when Fergason (1985) developed a system in which a nematic liquid crystal is dispersed in a water-borne polymer to form an emulsion. The liquid crystal becomes encapsulated by the polymer to form micrometre-size droplets. The emulsion is then sandwiched between two parallel plates and, depending on the final purpose of the product, the continuous polymer matrix could be dried or photo-cross-linked. In the literature this variation of a PDLC is called a "nematic curvilinearly aligned phase."

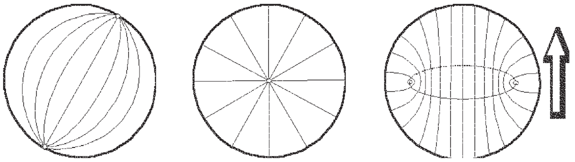


Figure 1
Nematic director alignment in spherical droplets. First shown is the droplet with planar director anchoring on its surface, leading to a pair of surface boojums on its poles. Radial anchoring results in a $(+1)$ point defect in the middle, a "hedgehog". Applying a strong electric field aligns most of the bulk of the droplet, splitting the hedgehog into a ring of $(+\frac{1}{2})$ disclination.

A more popularly used PDLC system was invented by Doane *et al.* (1986) at Kent State University in Ohio. In fabrication, a liquid crystal could be mixed with (i) a photopolymerizable reactive monomer, an example being the epoxy Epon 828 with an appropriate photocuring agent, (ii) a polymer melt, such as polymethylmethacrylate (PMMA) or polyvinylformal (PVF), or (iii) a mixture of polymer and solvent, compatible to the liquid crystal also, to form an initially homogeneous solution. The liquid crystal is then phase separated from the polymer matrix by one of the following methods: (i) polymerization-induced phase separation (PIPS); (ii) thermally induced phase separation (TIPS); or (iii) solvent-induced phase separation (SIPS). The advantage of this technology is that the phase-separation kinetics helps selection of a more uniform distribution of submicrometre droplet sizes.

A typical PDLC device, as shown in Fig. 2, is an electrically switchable composite film, consisting of a thin layer (10–30 μm) of micron-size dispersion of liquid crystal droplets in a solid polymer matrix, sandwiched between two transparent conductive glass or plastic supports. By matching the extraordinary refractive index of the liquid crystal in birefringent droplets, n_e, with that of the polymer binder, n_p, an

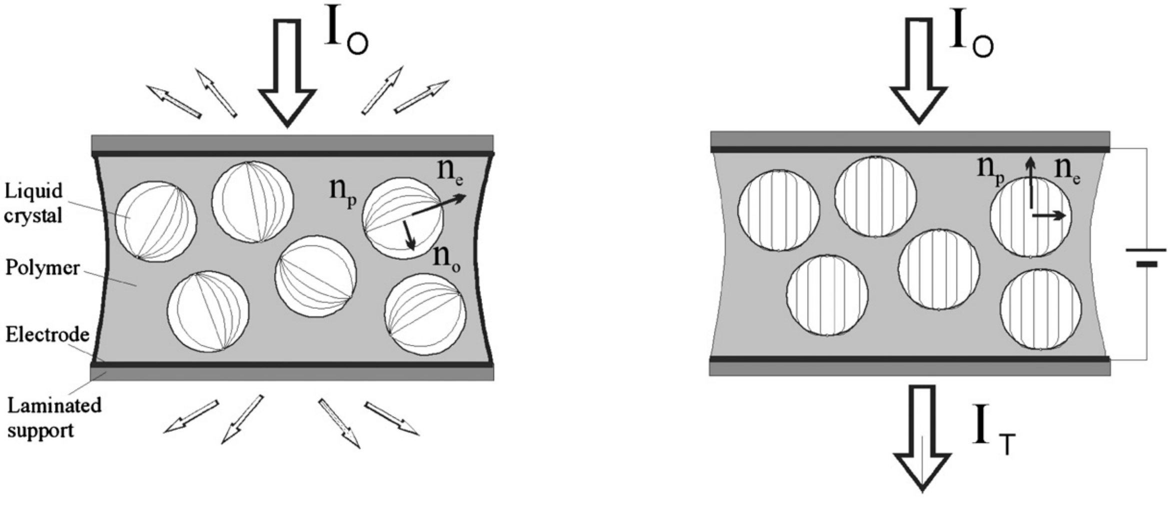


Figure 2
A scheme of PDLC device operation. Microdroplets are embedded in a polymer matrix, with transparent electrodes on both sides of the film. The refractive index of the polymer n_p is matched to the parallel (extraordinary) index n_e of the uniaxial birefringent droplets. In the OFF state all droplets are randomly oriented and strongly scatter the incident light I_o. In the ON state, the electric field aligns the director in each droplet, the refraction properties are matched with the matrix and the film becomes transparent to light.

electrically controllable light scattering medium transforms from a translucent, white appearance to a transparent appearance upon application of a voltage.

In the absence of fields (the OFF state), the PDLC film is highly scattering because the large variations of the refractive index ($\sim n_e$–n_o) occur on the scale of the wavelength of light: the material appears "opaque." On application of an electric field (the ON state), the liquid crystal molecules inside microdroplets are aligned with the field and PDLC film becomes "transparent" if the refractive indices are matched, $n_p = n_e$. The fast millisecond-scale switching response of director orientation inside droplets to the electric field causes a rapid interchange from "opacity" to "transparency" and is the basic operational principle of the PDLC device. Figure 3 shows a real-life device used, for instance, in architectural windows or privacy screens.

The physical principle of all liquid crystalline electro-optic devices is the effect of alignment of the nematic director (average direction of long molecular axes, the principal axis of birefringence) by the field. This dielectric coupling is local, with the characteristic energy determined by the uniaxial anisotropy of dielectric constant $\varepsilon_a = \varepsilon_{\|} - \varepsilon_{\perp}$, and in a droplet of radius R is of the order $\sim \frac{1}{2}\,\varepsilon_o \varepsilon_a E^2 R^3$. This is resisted either by the elastic energy stored in the director deformations, the Frank energy of liquid crystals, or by a viscous drag of rotating the bipolar director texture. The sketch in Fig. 2 shows both effects at play: the orientation of the boojum pair is along the applied field and the director profile inside each droplet is practically uniform.

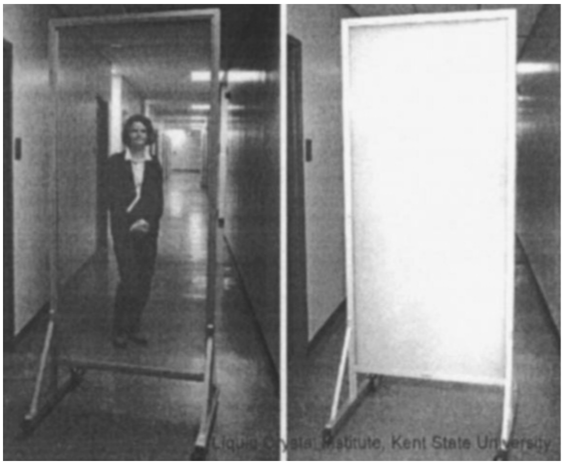


Figure 3
The Switching Window—a glass screen coated with 20 µm PDLC film which is electrically switching between the opaque OFF state and transparent ON state (image courtesy of LCI, Kent, Ohio).

In practice, if the liquid crystal droplets are symmetric, spherical, then one finds that the boojum pair is rather indifferent to its orientation and the rotation of the droplet axis can be achieved without an energy barrier, by a low voltage $V = E/d$ in a film of thickness d. In this case the principal resistance is viscous and the PDLC film will have a long response time: $\tau_{\mathrm{ON}} \approx \gamma_1 d^2 / \varepsilon_o \varepsilon_a V^2$ with γ_1 the nematic rotational viscosity, and an undetermined but very long recovery time τ_{OFF}. If, on the other hand, the droplets are asymmetric, i.e., deformed into an elliptical shape, there is a preferred orientation of the boojum pair. The electric field then has to work against the elastic resistance of the director field, which for spherical droplets of radius R is of the order of $\sim KR$. Thus, the electric field required for the switching of such a droplet has to exceed a threshold of the order of $E_s \approx (1/R)\sqrt{K/\varepsilon_o \varepsilon_a}$, or the applied voltage required to operate a PDLC film of thickness d is $V_s \geqslant (d/R)\sqrt{K/\varepsilon_o \varepsilon_a}$. This is $\sim(d/R)$ times higher than the characteristic Freedericks' voltage used to switch standard liquid crystal display cells; for a typical PDLC film of $d = 20\,\mu\mathrm{m}$ thick and droplets of under $R = 1\,\mu\mathrm{m}$, this may amount to a substantial increase and can be up to 200 volts. An interesting development by Drzaic (1988) and Wu *et al.* (1989) demonstrated the important role of asymmetry in the droplet shape in the relaxation dynamics. The recovery of the opaque OFF state proceeds with the relaxation time $\tau_{\mathrm{OFF}} \approx \gamma_1 R^2 / K(l^2 - 1)$ with l the aspect ratio of the droplet ellipsoid. For even weakly deformed droplets, $l = 1.1$, this relaxation time is $\sim 10\,\mathrm{ms}$, much faster than in conventional electro-optic nematic cells. By, for example, stretching the polymer film, one could produce a weakly elliptic shape of liquid crystal cavities consistently aligned along the stretching direction and thus remove the degeneracy in the position of the boojum pair on the droplet surface. This results in a recorded memory about the orientation in the OFF state and in a significant decrease in the device response times. Therefore, one finds an inevitable trade-off between the switching voltage and the relaxation times: decreasing one by altering the material parameters results in the undesirable increase of the other.

Another problem with the basic PDLC technology is that in the ON state, the material is transparent only for a light beam incident normally to the film and becomes hazy for obliquely incident light (where a difference between the matrix refractive index n_p and the ordinary index n_o of aligned droplets is being probed). An intriguing modification is possible if a liquid crystal polymer is used in PDLC instead of an isotropic binding matrix. The refractive indices n_{p1} and n_{p2} of the liquid crystal polymer could match those of the liquid crystal droplets. In the ON state, the liquid crystal polymer and the droplets are all aligned along the applied field; the refractive indices are matched for light incident at all angles and therefore the material is transparent in all directions. However, a different problem arises: in order for the refractive indices of the liquid crystal polymer and the liquid crystal to be matched, the mesogenic moieties of the liquid crystal polymer must have a structure similar to that of the liquid crystal. For such a liquid crystal and polymer system, phase separation becomes difficult.

The potential of composite technology in high-contrast electro-optic light shutters revitalized the general interest in microconfined liquid crystals in curved geometries. In order to solve the problem of refractive index matching, a new concept of dispersing a small amount of polymer in liquid crystals to form a PSLC composite material has been developed by Yang *et al.* (1992). In preparation, the liquid crystal is mixed with a reactive monomer, which is then polymerized, usually by photo-initiation. During polymerization, a surface alignment or external field can be used to control the orientation of the liquid crystal, and consequently the direction of the forming polymer network (see Fig. 4).

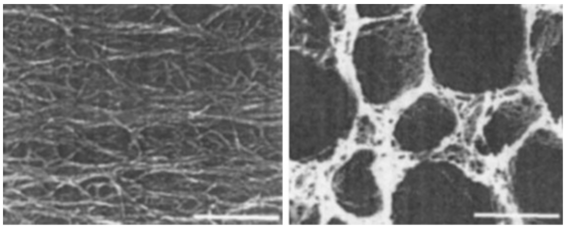


Figure 4
SEM images of a polymer fibre network formed in the pre-aligned nematic PSLC: cross-section views from the side (along the nematic director) and from the top (looking down the birefringence axis) confirm that the polymer strands are well aligned by the orientational field of liquid crystal. The white bar ($=5\,\mu$m) shows the length scales of polymer fibrils, their network and of polymer-free cavities (images courtesy of D.-K. Yang, LCI, Kent, OH).

In this way, once established, the polymer fibrils will have an aligning effect on the surrounding liquid crystal. For best results for such a mutual alignment the reactive monomer is selected to be mesogenic, such as 4,4′-bisacryloyloxy-1,1′-biphenylene (BAB), consisting of a biphenylene rigid core and the diacrylate reactive groups on both sides. The rigidity of the resulting polymer network is tuned by using analogous monomers but with flexible spacers, which give conformational freedom to the chains and a more rubbery nature to the composite.

Because of the low concentration of polymer, the refractive index matching problem does not arise in the PSLC composite. Particularly when the polymer is a mesogenic network, the scattering caused by the small refractive index mismatch between the liquid crystal and the polymer network is negligible. After polymerization, the polymer network acts in turn to control the orientation of the liquid crystal in the surrounding cavities. Now the aligning anchoring of the polymer network is a bulk effect and much more efficient than a surface alignment layer. The characteristic length scale of such a system is the cavity radius ξ, of the order of 1–2 μm. Because the director is anchored on the polymer strands, the system effectively responds to the electric field $E=V/\xi$, rather than the voltage. Assuming the polymer network is mechanically rigid and the director anchoring is strong, as with the BAB materials, the field required to generate a significant optical change is $E \geqslant (1/\xi)\sqrt{K/\varepsilon_o\varepsilon_a} \sim 10^6\,\mathrm{V\,m^{-1}}$ or higher.

A number of display applications have been developed, based on the polymer-stabilized twisted and super-twisted nematic, cholesteric and ferroelectric smectic C* liquid crystals. When the pitch of a

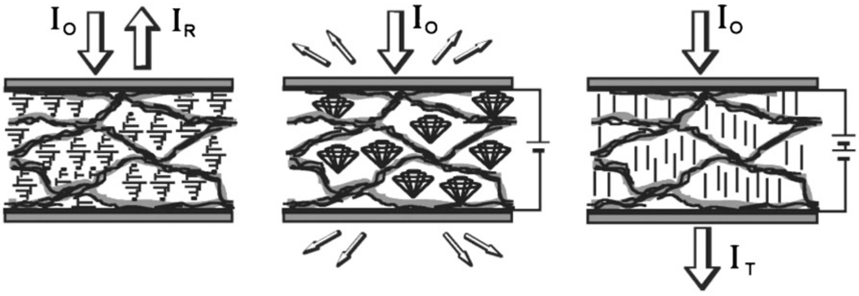


Figure 5
A scheme of cholesteric PSLC operation. In the OFF state, the polymer network follows the planar cholesteric helix, and preserves its memory in later stages by anchoring the director in appropriate configuration. The helical birefringent texture selectively reflects light. In the ON state, the director attempts to align along the field but is held back by the bulk anchoring on the polymer network—the result is a scattering conical texture. At very strong field the director is fully aligned and the PSLC composite becomes transparent.

cholesteric liquid crystal is in the region of ultraviolet to near-infrared light, the most prominent application is the bistable reflective cholesteric display. In such a composite, in the OFF state, the stabilizing polymer network follows the planar helical twisting and the material selectively reflects light (Bragg reflection from the periodic modulation of the refractive index).

In the ON state, at moderate voltage, the electric field transforms the director texture into a relatively random focal-conic arrangement determined by the competition between the field-aligning effect and the director anchoring on the polymer network strands. The resulting material is strongly scattering and has an opaque appearance. At a substantially higher field $E \geqslant (1/P)\sqrt{K/\varepsilon_o\varepsilon_a}$ with $P \sim 0.5\,\mu\text{m}$, the cholesteric helical pitch, the director can be fully aligned and make the composite transparent. However, the memory about the initial planar cholesteric alignment is permanently fixed in the polymer network and on switching the field off the optical properties quickly return to their original state. Such a sequence of electrically controlled optical states leads to a design of a variety of multimode light valves.

The randomly inhomogeneous PSLC composite with a dilute rigid polymer network should be contrasted with the liquid crystalline gel network, a strongly diluted example of a new class of materials—liquid crystalline rubbers and gels. Terentjev *et al.* (1999) have shown that when the growing and cross-linking polymer chains do not phase separate from a liquid crystal majority phase, the latter acts as an anisotropic solvent in a homogeneous gel with a weak rubber-elastic polymer network. In such a system the electric field causes not only the optical axis rotation, but also the mechanical deformation of the gel network. The resulting length scale (which determines the field required for electro-optic response) is determined by the ratio of Frank constant of the nematic liquid to the rubber modulus of the gel, $\xi \approx \sqrt{K/\mu}$, and can vary in a broad range from 0.01 to 5–10 μm, leading to a variety of possible field effects.

See also:

Main-chain Liquid Crystalline Polymers for Optical Applications#; Nematic Liquid Crystalline Polymers: Defects and Dislocations#; Nematic Liquid Crystals: Defects#; Cholesteric Liquid Crystals: Optical Properties#; Cholesteric Liquid Crystals: Distortion of Structure by External Fields#; Cholesteric Liquid Crystals: Defects#

Bibliography

Bahadur B (ed.) 1990*Liquid Crystals: Applications and Uses.* World Scientific, Singapore, Vol. 1

Crawford G P, Zumer S (eds.) 1996 *Liquid Crystals in Complex Geometries.* Taylor and Francis, Philadelphia, PA

De Gennes P G, Prost J (eds.) 1993 *The Physics of Liquid Crystals.* Oxford University Press, Oxford

Doane J W, Vaz N S, Wu B G, Zumer S 1986 Field controlled light scattering from nematic microdroplets. *Appl. Phys. Lett.* **48**, 269–71

Drzaic P S 1988 Reorientation dynamics of polymer dispersed nematic liquid crystal films. *Liq. Cryst.* **3**, 1543–59

Fergason J L 1985 Polymer encapsulated nematic liquid crystals for display and light control applications. *Soc. Inf. Disp. Dig.* **XVI**, 68–9

Terentjev E M, Warner M, Meyer R B, Yamamoto J 1999 Electro-mechanical Fredericks effects in nematic gels. *Phys. Rev. E* **60**, 1872–9

Wu B G, Erdmann J H, Doanne J W 1989 Response times and voltages for PDLC light shutters. *Liq. Cryst* **5**, 1453–65

Yang D-K, Chien L -C, Doane J W 1992 Cholesteric liquid crystal–polymer dispersions for haze-free light shutters. *J. Appl. Phys.* **60**, 3102–4

Zurek W H 1996 Cosmological experiments in condensed matter physics. *Phys. Rep.* **276**, 177–221

E. Terentjev
University of Cambridge, UK

Lumber: Laminated Veneer

Laminated veneer lumber (LVL), also known as parallel-laminated veneer, is one of the most suitable products for the dispersion of natural defects of wood (knots, etc.) by glue lamination of veneer. LVL can be produced with higher yields, in a greater range of dimensions and shapes, and with less time and labor than glued laminated timber or plywood. LVL is composed of mostly parallel grain laminations of veneer bonded to obtain lumber-size thickness. LVL has been approved as an engineering material with reliable strength and stiffness.

Most LVL is produced in flat shapes. When thin veneers are used as laminas, the manufacture of molded LVL with a small radius of curvature is possible. Curved (molded) LVL is used for door or window fittings and for furniture frameworks. Even cylindrical LVL with hollow or reinforced cores has been produced (Sasaki *et al.* 1996).

Development of LVL began with the production of high-strength wooden members in aircraft in the 1940s. In the following decades the machinability, natural appearance, and uniformity of mechanical properties of LVL were used to advantage in the production of curved furniture parts.

Since the supply of high-quality saw logs has diminished, LVL has become more important because of its higher yield potential, adaptability to engineering end-use designs, and the introduction of automatic production methods. Since the beginning of the 1970s LVL has also been used for structural members (e.g., tension chords of trusses or I-beams and outer laminas of glued laminated beams) in place of lumber components because of its reliability in strength.

1. Processing

LVL is manufactured in a manner similar to plywood but contains mostly parallel laminas. Rotary-cut veneers are usually used. With the decrease in the diameter of available logs, more efficient lathes have been developed to get higher veneer yields. Rotary lathes with peripheral driving devices make it possible to peel logs down to very small (50mm) core diameters.

Veneers are sometimes treated with preservatives, fire retardants, or other chemicals before drying. These chemicals can also be mixed with adhesive. Normally, veneers are dried to about 5% moisture content in conventional hot air circulating, jet, or press dryers. Thin veneers can be dried more economically than thick veneers; however, the production of LVL with thinner veneer requires more adhesive. The balance of these factors, as well as the improved properties of LVL achieved with thinner veneer, must be considered in the choice of veneer thickness.

Phenolic-type or other equivalent adhesives (e.g., melamine, isocyanates) are used in the production of structural LVL, while urea resin is generally used for furniture or nonstructural materials.

Butt joints of veneer ends are significant defects in LVL. The joints must be staggered and distributed as evenly as possible. When the number of laminations is high, butt joints are acceptable because the decrease in the mechanical properties becomes less. Scarf joints in veneer ends are also employed in some cases. Although LVL with scarf-jointed veneer has better strength and appearance, the process is more complex than for butt joints. To obtain high tensile strength in a simpler process, crushed lap joints have been used in commercial LVL (Micro-Lam), where high processing pressure is applied to crush overlapping veneer ends.

Since LVL production started as an extension or variation of plywood production, most LVL is hot pressed in multiopening presses, where LVL is limited in length. Modern plants for the production of structural LVL tend to employ a continuous veneer assembly and hot pressing, in which the joints are well distributed. This makes LVL production less labor intensive than the production of plywood or glued laminated timber. A tractor-tread press with electrical heating units is employed to produce high-tensile-strength LVL (Micro-Lam). Another type of tractor-tread press uses radio frequency heating in which electrodes are set at the surfaces of the treads. Steel belt type presses use residual heat from veneer drying or hot platens set behind the belts. With these continuous presses, high yield and highly efficient production of continuous LVL is feasible. In the case of molded LVL, layers of thin veneer are pressed between male and female molds with radio frequency heating devices.

Cylindrical LVL is produced using a spiral winding method (Fig. 1) (Yamauchi *et al.* 1998). For instance, to manufacture an LVL cylinder with a 500 mm diameter, a continuous glued-veneer tape of 150 mm width (parallel to the fiber direction) and 2.5 mm thickness is wound spirally on a steel mandrel. A pressure of about 0.5 MPa is applied with an elastic belt. The steel mandrel, formed veneer cylinder, and pressure belt are then heated in an oven to cure the adhesive.

2. Mechanical Properties

Figure 2 shows how the variations of strength are reduced in LVL compared to solid lumber from the same logs. The reliability in strength increases with increasing number of laminas. The relative improvement in strength is higher when the quality of the logs is lower. Similar relationships exist between the stiffness of LVL and solid lumber.

The allowable design stress f_a can be estimated by the approximate form:

$$f_a = \frac{f_m - \eta\sigma}{2.1}$$

where f_m is the average strength, σ is the standard deviation, η is a coefficient depending on the shape of the strength distribution, and the factor 2.1 includes the effects of long-term loading, a safety factor, etc. In general, the 5% exclusion limit f_5 (the stress value below which the probability of failure occurring is 5%) is taken to be ($f_m - \eta\sigma$). If the distribution is Gaussian, η is 1.645. A number of experiments with LVL have shown that η should be taken as 2.0, considering safety factors in use (Koch 1973, Bodig and Jayne 1982).

As can be seen in Fig. 2, the estimated allowable stress of LVL will be far greater than that of solid lumber. LVL manufactured commercially from 2.5–3.0 mm thick Douglas fir veneer (Micro-Lam) with 13 or more plies and crushed lap joints has been approved officially as a material having an allowable stress of 15.09 $\mathrm{MN\,m^{-2}}$ in tension and 19.21 $\mathrm{MN\,m^{-2}}$ in bending. The strength reduction due to butt joints in LVL decreases with increasing number of plies. The reduction is very small in 9-ply products and negligible for 12 or more plies. As derived from fracture mechanics theory, the tensile strength of LVL is inversely proportional to the square root of the thickness of the laminas. The distance between the butt joints in adjacent layers should be more than 16 times the lamina thickness to prevent strength reduction due to stress interaction (Laufenberg 1983).

The tensile strength of products produced by end-jointing short LVL, in general, is more than 90% of

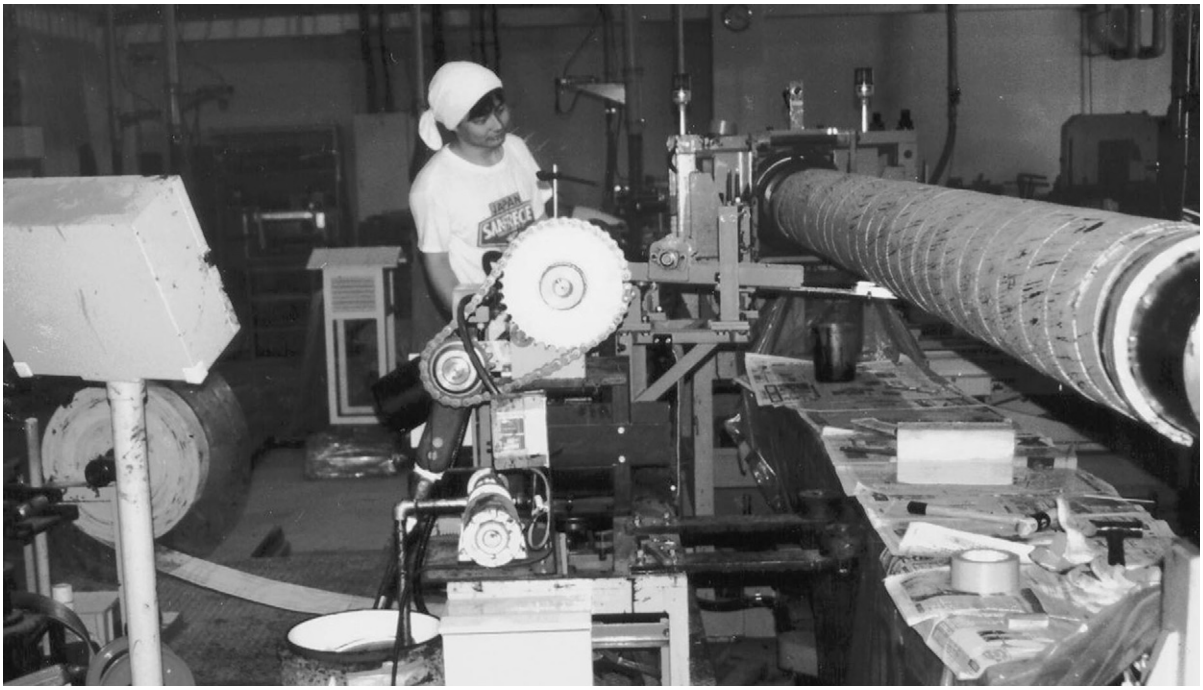


Figure 1
Cylindrical LVL manufacture using a prototype spiral winder.

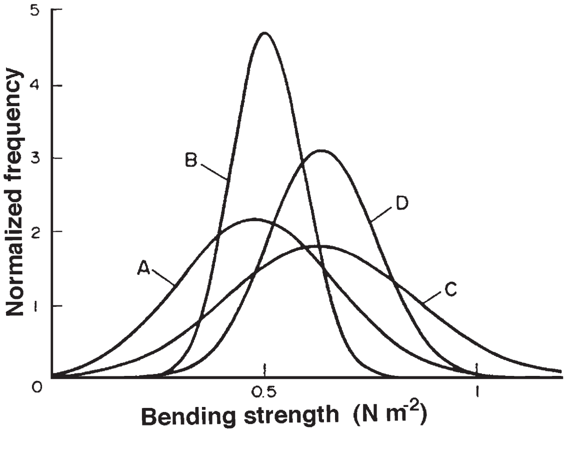


Figure 2
Comparison of bending strength between solid lumber and LVL processed from the same logs. A: solid sawn, 2″ × 4″, Siberian larch; B: 8-ply LVL, same log as A; C: solid sawn, 2″ × 4″, southern pine; D: 6-ply LVL, same log as C.

the strength of unjointed LVL for scarf joints with a 1:9 slope, more than 70% for finger joints of 27.5mm in length, and more than 60% for minifinger joints of 10mm in length.

Low strength perpendicular to the long axis (i.e., cleavage and shear in the plane parallel to the fibers) is a drawback of LVL. This arises from lathe checks in veneers created during peeling. However, the average strength increases and the variation decreases with increasing number of plies, and thus the allowable stress perpendicular to the grain can be estimated to be the same value as for solid lumber.

3. *Grading and Quality Control*

One of the most promising characteristics of LVL is the capability of ensuring its structural performance by grading veneer according to quality and selecting the optimum combination or placement of the veneer. Visual grading of veneer which has been widely used in the plywood industry is not totally satisfactory for screening veneer used in stress-rated LVL. Mechanical methods of testing modulus of elasticity of veneer in tension are used in research, but the commercial application of this is not feasible. The stress-wave timing technique is one of the most promising nondestructive methods of estimating properties of each veneer piece before it is laminated into LVL in the industrial process (Jung 1982).

4. *Applications*

The furniture and fittings industries prefer LVL because of its ease of processing, uniform properties, good adaptability to curved parts with small radius, stability, and better edge appearance. Arched door frames, staircases, chairs, beds, cabinets, counter tables, wardrobes, etc., are popular uses of this material.

The uniform mechanical properties of LVL, and especially its high allowable stress in tension compared to its weight, make this material useful for flanges of I-beams combined with plywood webs or steel pipe lattice webs. Such I-beams are widely used as joists in house construction. Open-web trusses combining LVL and steel pipe are used as roof trusses and second floor joists in large buildings such as factories and warehouses. These open-web trusses are also used for arched roofs of very large structures such as a 120m span football stadium and exhibition halls. The uses are extending to stringers and decks of highway bridges, stringers of trailer cars, cross-ties, scaffold planks, and to other engineered members in wood construction (Kunesh 1978). Cylindrical LVL can be used for columns in heavy wooden constructions and as arched frames with special functions such as ducting for ventilation, electrical wires, water pipes, etc. (Sasaki and Kawai 1998).

See Also: Wood: Preservative Treated#; Wood: Strength and Stiffness*; Wood: Structural Panel Processes*; Wooden Structures#

Bibliography

Bodig J , Jayne B A 1982 *Mechanics of Wood and Wood Composites*. Van Nostrand Reinhold, New York

Jung J 1982 Properties of parallel-laminated veneer from stress-wave-tested veneers. *For. Prod. J.* **32**(7), 30–5

Koch P 1973 Structural lumber laminated from 1/4-inch rotary-peeled southern pine veneer. *For. Prod. J.* **23** (7), 17–25

Koch P 1985 *Utilization of Hardwoods Growing on Southern Pine Sites*, Agriculture Handbook no. 605. US Government Printing Office, Washington DC, 3

Kunesh R H 1978 Micro-Lam: structural laminated veneer lumber. *For. Prod. J.* **28**(7), 41–4

Laufenberg T L 1983 Parallel-laminated veneer: processing and performance research review. *For. Prod. J.* **33** (9), 21–8

Sasaki H, Yamauchi H, Koizumi A, Iijima Y, Tamura Y, Sasaki T 1996 Cylindrical LVL for structural use made by spiral-winding method. *Proc. Int. Wood Engineering Conf.* Omnipress, WI, Vol. 3 538-40

Sasaki H , Kawai S 1998 Development of conversion technologies of low-grade wood resources into higher-grade laminated veneer products. *Proc. Workshop 40th Anniversary of Chinese Academy of Forestry*. China Agricultural Scientech Press

Yamauchi H, Miura I, Sasaki T, Koizumi A, Kawai S, Sasaki H 1998 Manufacturing condition and mechanical properties of cylindrical LVL manufactured by spiral-winding method. *J. Soc. Mater. Sci. Jpn.* **47**(4), 350–5

H. Sasaki
Akita Prefectural University, Akita, Japan

M

Magnets: Bonded Permanent Magnets

Bonded magnets are generally prepared by blending coercive powder with a binder followed by compacting or molding the material to the final shape. Because of the presence of the binder the energy product of bonded magnets is always lower than that of their sintered counterparts. However, two great advantages of bonded magnets are their superior mechanical properties and the fact that the magnets can be prepared under net shape conditions. Bonded and sintered magnets of ferrites have been known for many decades. Both types of magnets have very low energy products when compared to sintered rare earth-based magnets. The excellent properties of rare earth-based magnets are, however, accompanied by a high price.

The last decade of the twentieth century saw an increased use of actuators and small electromotors not only of low weight but also of low energy consumption, as in battery-driven motors in particular. There has also been an increased use of electronically controlled, electrically actuated systems in automobiles, especially for critical vehicle control functions. All this and the increased use of electromagnetic devices in the electronic and electric industry has led to an enhanced pressure to find improved or alternative permanent magnet materials lying between the low-cost low-performance ferrite magnets and the high-cost high-performance sintered NdFeB magnets. Bonded NdFeB magnets meet most of the criteria mentioned and their use is growing steadily. For this reason bonded types of rare earth permanent magnets will be discussed below in somewhat more detail.

1. *Bonding Technologies*

Bonded magnets are generally prepared by blending coercive powder with a binder or by encapsulating coercive powder in a binder, followed by compacting or molding the material to the final shape. The powder of the magnetic materials used in bonded magnets may be ferrites, alnico, samarium cobalt or neodymium–iron–boron. In some cases also hybrid powders are used.

Depending on the binder, the magnets produced may be flexible or rigid. Nitrile rubber or vinyl are commonly used for the former type whereas for the latter type nylon, teflon, polyester, or thermoset epoxies are employed.

There are four main manufacturing routes for bonded magnets: calendering, injection molding, extrusion, and compression bonding.

(i) Calendering involves processing the material between rollers and leads to continuous strips of adjustable thickness and length. This process is schematically shown in Fig. 1. The material used for calendering is usually a mixture of the magnetic powder and a thermoplastic binder. This mixture is processed between rollers forming a continuous strip that may be up to 100 m in length. Magnetic loading can be up to 70% by volume.

(ii) The injection molding process is schematically represented in Fig. 2. This process involves forcing the heated mixture of magnetic powder and thermoplastic binder via tubes into a mold where it is allowed to cool and harden. In this case too, magnetic loading can reach 70%, especially when using powder of spherical shape (Ma *et al.* 2002).

(iii) Extrusion makes use of an orifice through which the compounded material is squeezed, as illustrated in Fig. 3. During the extrusion process the orifice is heated and the profile is controlled as the compounded material cools. Depending on the binder, the cooled product can be firm, as for flexible end products, or rigid if a nonflexible end product is desired. The loading can reach slightly higher values (up to 75% by volume) than in the two previous cases.

(iv) Compression bonding uses magnetic powder blended with an epoxy that has been liquefied by dissolution in an organic solvent. The latter is evaporated during the blending process, leaving the magnetic particles in an epoxy-encapsulated state. After

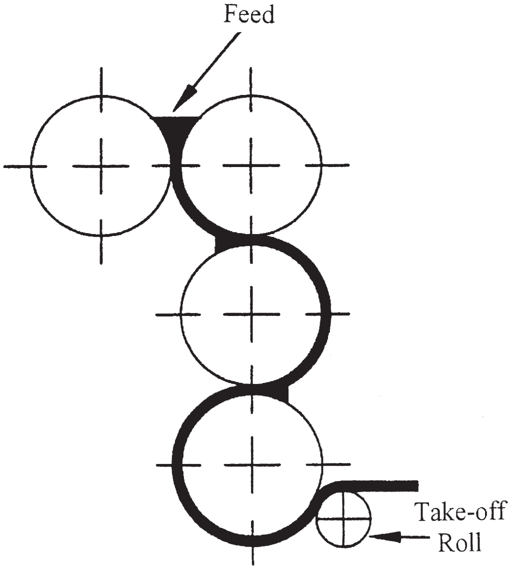


Figure 1
Schematic diagram showing the manufacturing of magnetic strips by calendering (after Ormerod and Constantinides 1997).

---

* Cross references marked by an asterisk are included in this volume.

$^{\#}$ Cross references marked by a hash can be found by consulting the Encyclopedia of Materials: Science and Technology.

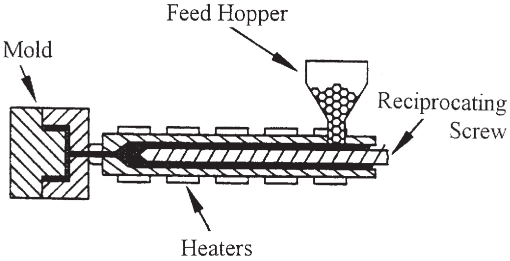


Figure 2
Schematic representation of the injection molding process. This process involves forcing the heated mixture of magnetic powder and thermoplastic binder via tubes into a mold where it is allowed to cool and harden (after Ormerod and Constantinides 1997).

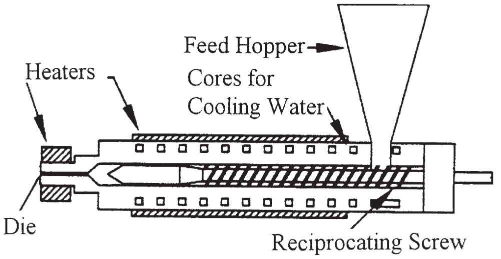


Figure 3
Schematic representation of the extrusion process. The compounded magnetic material is squeezed through an orifice. During the extrusion process the orifice is heated and the profile is controlled as the compounded material cools (after Ormerod and Constantinides 1997).

drying, the coated powder is flowed into a conventional powder press (Fig. 4) and compacted into the desired shape. Finally, the compacted magnet is cured in an oven at temperatures of around 150–175 °C. This process can reach magnetic loadings up to 80%.

Because of the presence of the binder, the magnetic remanence is always lower than that of the fully dense magnetic material, as obtained, for instance, by sintering. The situation is even less favorable for the energy product, being roughly proportional to the remanence squared. There are, however, obvious advantages. The major advantage is that bonded magnets can be produced in net shape or near-net shape. Even fairly complicated shapes may be involved, as in injection molding. Another advantage is their superior mechanical properties. Bonded magnets are mainly manufactured from ferrite powder or coercive NdFeB-based powders. Bonded magnets based on NdFeB, owing to their superior intrinsic magnetic properties, offer substantial advantages in terms of size, weight, and performance over bonded ferrite magnets and even over sintered ferrite magnets.

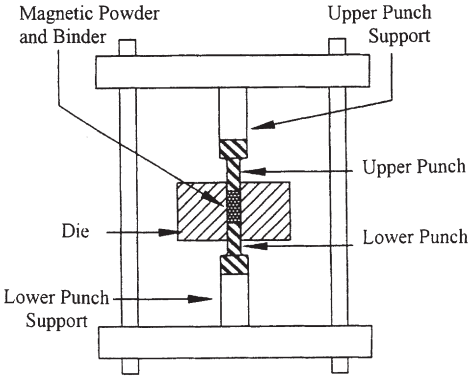


Figure 4
Diagram showing the compression bonding process. The magnetic powder is blended with an epoxy that has been dissolved in an organic solvent. The latter is evaporated during the blending process, leaving the magnetic particles in an epoxy-encapsulated state. The coated powder is flowed into a conventional powder press and compacted into the desired shape. The compacted magnet is finally cured in an oven (after Ormerod and Constantinides 1997).

2. The Permanent Magnet Materials

The most common types of magnets used at present are hard ferrite magnets, rare earth-based magnets such as SmCo or NdFeB, and alnico-type magnets. Of these the alnico magnets have only a modest coercivity which leads to nonlinear demagnetization characteristics. For this reason their applicability is very limited compared with the other two types. The hard ferrites have higher coercivities than the alnico magnets and their demagnetizing characteristics are linear. However, the remanence and the concomitant maximum energy product are already low and are further decreased by bonding. Because of their low cost, ferrite magnets are still widely applied, although most of the corresponding magnetic devices are rather bulky and often give far from optimal performance. Ferrite permanent magnets currently dominate automotive and many other applications due to their low cost and proven long-term stability (see *Alnicos and Hexaferrites*[#]).

The rare earth based magnets have high values of coercivity which gives them linear demagnetization characteristics (see *Rare Earth Magnets: Materials*[#]). They have high remanences and typical values of the energy products reached in sintered magnets are 150 kJ m^{-3} for $SmCo_5$, and 300 kJ m^{-3} for $Nd_2Fe_{14}B$. The first types of magnets are frequently used in high-temperature applications, which possibility is lost in bonded magnets (see *Magnets: High-temperature*[#]). $SmCo_5$ type magnets are expensive owing to the high price of both samarium and cobalt.

The situation is more favorable for NdFeB magnets because neodymium is cheaper than samarium and iron is much cheaper than cobalt, the powder metallurgical processing arts being comparable with those of $SmCo_5$. Hence the performance/price ratio for $Nd_2Fe_{14}B$ is better than for $SmCo_5$. For this reason the market for sintered $Nd_2Fe_{14}B$ magnets has flourished and is still growing (see *Magnets: Sintered*#).

Although bonded magnets can be manufactured from all of the materials mentioned above, only bonded ferrite magnets and bonded $Nd_2Fe_{14}B$ magnets have penetrated into the market to an appreciable extent. The ferrites can be easily obtained in powder form (see *Alnicos and Hexaferrites*#). A somewhat special processing of the ferrite powders leads to hexagonal platelets that can be easily aligned mechanically during the formation process of flexible bonded magnets. The situation is more difficult in the case of $Nd_2Fe_{14}B$, because a simple powder metallurgical route from cast or annealed $Nd_2Fe_{14}B$ ingots does not generally lead to powders of sufficiently high coercivity for use in bonded magnets. Coercive NdFeB powders can be obtained though by melt spinning (see *Metallic Filaments**). In this technique a fine stream of molten alloy is sprayed onto the outer surface of a rapidly spinning wheel, leading to thin rapidly quenched ribbons or flakes. During the melt spinning the material is protected from oxidation by performing this process in a protective atmosphere or in vacuum. The quenching rate is of the order of 10^5 K s^{-1} and can be varied by changing the rotational speed of the spinning wheel. Different quenching rates lead to different microstructures which in turn determine the magnetic properties of the melt spun material. Optimum melt spinning conditions lead to a nanocrystalline alloy consisting of fine grains (typically 30 nm) of the compound $Nd_2Fe_{14}B$, surrounded by a thin layer of a neodymium-rich eutectic phase. In practice, more reliable and reproducible results can be obtained by using a slightly higher quenching rate and subsequently annealing the melt spun material under carefully controlled conditions. The melt spun material is fairly brittle and can be ground to a fine powder suitable for manufacturing bonded magnets. Because the $Nd_2Fe_{14}B$ grains have a random orientation, these bonded magnets are isotropic. Powder particles of spherical shape flow better in the injection molding process, which allows a higher loading factor. Such powder can be prepared by an inert gas atomization process with typical values for the mean particle diameter of 45 mm. An additional advantage of the atomization process is the high production rate and low processing costs of the powders (Ma *et al.* 2002).

A different route leading to coercive NdFeB powders consists of the so-called HDDR process (see *Magnets: HDDR Processed*#). This process involves essentially four steps: **h**ydrogenation of $Nd_2Fe_{14}B$ at low temperatures, **d**ecomposition of $Nd_2Fe_{14}BH_x$ into $NdH_{2.7}$ + Fe + Fe_2B, **d**esorption of H_2 gas from $NdH_{2.7}$, and **r**ecombination of Nd + Fe + Fe_2B into $Nd_2Fe_{14}B$. This process profits from the fact that the formation of $Nd_2Fe_{14}B$ grains in the last step is a solid-state reaction and hence proceeds at a rate considerably lower than during solidification from the melt during a normal casting process. The average $Nd_2Fe_{14}B$ grain size remains in the nanometer range and gives rise to sufficiently large coercivities.

A further advantage is the fact that the HDDR process can successfully be used to obtain anisotropic particles. Takeshita and Nakayama (1992) discovered that additives of zirconium, hafnium, and gallium, in particular, are very effective in producing anisotropic HDDR powder. The amount of additive required is surprisingly small (for instance, $Nd_{12.5}Fe_{69.9}Co_{11.5}B_6Zr_{0.1}$). Microscopic investigations described by Harris (1992) revealed that large facetted HDDR grains had formed within the original as-cast grain of the alloy. These facetted grains have a common orientation, which is probably the same as that of the original grain. The anisotropic nature of HDDR powders of alloys like $Nd_{12.5}Fe_{75.9}Co_{11.5}B_8Zr_{0.1}$ can thus be visualized by assuming that the HDDR grains have nucleated and grown within an original as-cast grain region from submicron grains, the latter having a common orientation (Harris 1992). Evidently the effect of the additive is to bring about nucleation centers for nucleation and growth of the HDDR grains, the latter grains having kept the orientation of the original cast grain.

Tomida *et al.* (1996) have used x-ray diffraction to establish a correlation between the anisotropic nature of the ultimate HDDR powder and the amount of $Nd_2Fe_{14}B$ phase remaining unreacted in the hydrogenation process. TEM studies made by Tomida *et al.* on powder hydrogenated under optimum energy product conditions showed that after hydrogenation the powder consists mainly of coarse-grained α-Fe and Fe_2B, with nanocrystalline particles embedded in between. These particles were identified by electron diffraction as NdH_2 particles. However, many of the particles were identified as $Nd_2Fe_{14}B$ particles having a crystallographical orientation almost the same as that of the original cast $Nd_2Fe_{14}B$ grains. Energy-dispersive spectra furthermore showed that these particles are of higher cobalt and gallium concentration than corresponding to the average concentration of the starting alloy. These results have led Tomida *et al.* to propose that this type of nanocrystalline $Nd_2Fe_{14}B$ particles serve as initiation centers in the recombination process and are the origin of the orientational memory effect in HDDR powders.

A different category of materials that appears to be promising is nanocrystalline rare earth-based composite magnets. Under special circumstances, two-phase composite materials can demonstrate a most interesting coercivity behavior. Such a behavior has been described by Kneller and Hawig (1991), who investigated the combined effect of two finely dispersed and mutually exchange-coupled magnetic phases. One of these phases has a large uniaxial anisotropy constant and is able to generate a high coercivity. By contrast, the second phase is magnetically soft. It has a larger magnetic ordering temperature and concomitantly a larger average exchange energy than the hard phase. It is the comparatively high saturation magnetization of the soft phase that, when the latter is exchange coupled to the hard phase, provides a high remanence to the composite magnet. The possibility to prepare magnets showing remanence enhancement has triggered extensive research in this area (see *Magnets: Remanence-enhanced*[#]).

In most of the systems for which remanence enhancement has been reported, the magnetically soft phase is α-Fe or an iron-rich or cobalt-rich alloy. Examples of magnetically hard phases are $Nd_2Fe_{14}B$, $Sm_2Fe_{17}N_3$, Sm_2Co_{17}, and $Nd(Fe,Mo)_{12}N_x$. The microstructures of all these composite magnets have in common that they consist of a very fine distribution of the magnetic particles, falling into the nanometer range. In order to reach this fine distribution various techniques are employed, including melt spinning and mechanical alloying (see *Magnets: Mechanically Alloyed*[#]). This group of materials is referred to as lean rare earth permanent magnets. Their advantages compared with the standard alloys are their excellent corrosion resistance and the fact that they reach saturation in a comparatively low applied field. A disadvantage is their relatively low coercivity. The possibility of using these materials in resin-bonded magnets has been described by Croat (1997).

Another interesting group of materials are interstitially modified R_2Fe_{17} compounds. Although the low Curie temperatures and comparatively low magnetocrystalline anisotropies make the R_2Fe_{17} compounds less attractive for applications as permanent magnet materials, considerable improvements with respect to Curie temperature anisotropy and coercivity have been reached by forming interstitial solid solutions obtained by combining these materials with carbon or nitrogen. The composition of the corresponding ternary nitrides and carbides $R_2Fe_{17}C_x$ and $R_2Fe_{17}N_x$ is generally believed to be restricted to the range $0 \leq x \leq 3$. More details about the formation ranges and location of the interstitial atoms the lattice are described in the review of Fujii and Sun (1995).

Resin-bonded magnets from nitrogenited Sm_2Fe_{17} powders have been prepared with $BH_{\max} = 136\,\mathrm{kJ\,m^{-3}}$, $B_r = 9.0\mathrm{T}$, and $\mu_{oB}H_c = 6.5\mathrm{T}$. In order to explore the favorably low temperature coefficient of the coercivity in magnet bodies suitable for high-temperature applications Rodewald *et al.* (1993) and Kuhrt *et al.* (1993) have investigated tin- and zinc-bonded magnets. In these cases, however, the obtained remanences were fairly low ($B_r < 0.7$ T).

3. Applications of Bonded Magnets

Bonded ferrite magnets are commercially available in isotropopic and anisotropic grades (see Table 1). The calendered material is commonly called sheet. Sheet magnets are mainly used by screenprinters for advertising specialties, novelties, and picture frames. Sign manufacturers use it for vehicle signs, but there are also medical applications such as needle counters. Extruded profile flexible magnets are called strip. They are used for a variety of gasket applications such as for refrigerator doors, shower doors, and removable storm windows. Other applications are in small motors and stereo speakers. Injection-molded ferrite magnets are mainly used in automotive applications, office automation, and in small motors. A great advantage of the ferrite magnets is that they are inexpensive and the raw materials are readily available.

Bonded SmCo-type magnets are used in applications where the temperature coefficients of remanence and coercivity have to be low and where a high corrosion resistance is required. Compared to bonded ferrite and NdFeB magnets their market share is low because of the high price of the raw materials.

Table 1
Maximum energy products (kJ m^{-3}) of commercially available bonded magnets (from Ormerod and Constantinides 1997).

Magnet powder	Manufacturing route	Magnetic isotropic	Magnetic anisotropic
Ferrite	Calendered	4.0–5.0	11.2–12.8
	Extruded	3.2–4.8	9.6–12.0
	Injection molded	4.0–6.4	12.0–13.6
SmCo	Injection molded		68–76
	Compression		104–136
NdFeB	Calendered	39–41	
	Injection molded	40–42	76–88
	Compression	64–88	112–128

The majority of bonded NdFeB magnets are used in personal computers but they are also used in a fairly wide range of automotive, consumer, and office products including VCRs, camcorders, and printers, and in smaller appliances such as watches, clocks, timer switches, and cameras. Most of the bonded NdFeB magnets go into various types of electric motors. The most prominent types of electric motors are spindle motors and stepper motors. Stepper motors have fairly small sizes and are characterized by their ability for precise incremental position or speed adjustment with electric pulses from digital controllers. In modern stepper motors, the stators consist of the coils system and the rotors are made of rare earth permanent magnets. The advantage of using rare earth permanent magnets is their high coercivity, which makes it possible to axially magnetize the rotor through the thickness with a large number of poles on the circumference. This allows fairly low step angles. The number of spindle motors produced is lower than the number of stepper motors. Spindle motors are much larger than stepper motors. This is the reason why they still consume a relatively large percentage of bonded NdFeB magnets. The largest application of electric motors is in computer peripherals, where they serve, for instance, as hard disk drive spindle motors, CD-ROM pick-up motors, CD-ROM spindle motors, floppy disk drive spindle motors, and floppy disk head actuators.

For a more detailed discussion of bonded magnet manufacturing, their performance, and their applications the reader is referred to the review articles of Croat (1997), Ormerod and Constantinides (1997), and Ma *et al.* (2002).

See also: Magnetic Materials: Domestic Applications#

Bibliography

Buschow K H J 1997 Magnetism and processing of permanent magnet materials. In: Buschow K H J (ed.) *Magnetic Materials*. Vol. 10. Elsevier Science, Amsterdam, Chap. 4

Croat J J 1997 Current status and future outlook for bonded neodymium permanent magnets. *J. Appl. Phys.* **81**, 4804–9

Fujii H, Sun H 1995 Interstitially modified intermetallics of rare earth and 3d elements. In: Buschow K H J (ed.) *Handbook of Magnetic Materials*. Vol. 9. Elsevier Science, Amsterdam, Chap. 3

Harris I R 1992 The use of hydrogen in the production of Nd-Fe-B-type magnets and in the assessment of Nd-Fe-B-type alloys and permanent magnets—an update. *Proc. 12th Int. Workshop Rare Earth Magnets and their Applications*, Canberra, 1992, pp. 347–71

Kneller E F, Hawig R 1991 The exchange spring magnet: a new material principle for permanent magnets. *IEEE Trans. Magn.* **27**, 3588–600

Kobayashi K, Iriyama T, Imaoka N, Suzuki T, Kato H, Nakagawa Y 1992 Permanent magnet material of $Sm_2Fe_{17}N_x$ compounds obtained by reaction in NH_3-H_2 mixture gas atmosphere. *Proc. 12th Int. Workshop Rare Earth Magnets and their Applications*, Canberra, 1992, 32–43

Kuhrt C, Schnitzke K, Schultz L 1993 Development of coercivity in $Sm_2Fe_{17}(N,C)_x$ magnets by mechanical alloying, solid–gas reaction, and pressure-assisted zinc bonding. *J. Appl. Phys.* **73**, 6026–8

Ma B M, Herchenroeder J W, Smith B, Suda M, Brown D N, Chen Z 2002 Recent developments in bonded NdFeB magnets. *J. Magn. Magn. Mater.* **239**, 418–23

Ormerod J, Constantinides S 1997 Bonded permanent magnets: current status and future opportunities. *J. Appl. Phys.* **81**, 4817–20

Rodewald W, Wal B, Katter M, Veliscescu M, Schrey P 1993 Microstructure and magnetic properties of Zn- or Sn-bonded $Sm_2Fe_{17}N_x$ magnets. *J. Appl. Phys.* **73**, 5899–901

Takeshita T, Nakayama R 1992 Magnetic properties and microstructures of the Nd-Fe-B magnet powders produced by HDDR process. *Proc. 12th Int. Workshop Rare Earth Magnets and their Applications*, Canberra, 1992, pp. 670–81

Tomida T, Choi P, Maehara Y, Uehara M, Tomizawa H, Hirosawa H 1996 Origin of magnetic anisotropy formation in the HDDR-process of $Nd_2Fe_{14}B$ based alloys. *J. Alloys Compds.* **242**, 129–35

K. H. J. Buschow
University of Amsterdam, The Netherlands

Materials Selection and Mechanical Design

1. Viability

The viability of a new material in a given application depends on the balance between its performance and its cost. There are three steps in evaluating it. The first is the technical assessment (Fig. 1, upper circle). Performance metrics are identified and evaluated for competing solutions for the design. Each application can be modeled in this way in order to provide a basis for performance comparisons between new material solutions and existing solutions. The second step is the analysis of cost (Fig. 1, lower left-hand circle): how much does it cost to achieve a given performance metric? The quantity of material required to meet constraints on stiffness, on strength, on energy absorption, etc., is calculated from straightforward technical models. The cost, *C*, of producing this quantity of material in the desired shape is the output of the cost model, as described in Sect. 3. The final step is that of assessing value (Fig. 1, lower right-hand circle): is the change in performance worth the change in cost? Balancing performance against cost is an example of multiobjective optimization. It is discussed in Sect. 4.

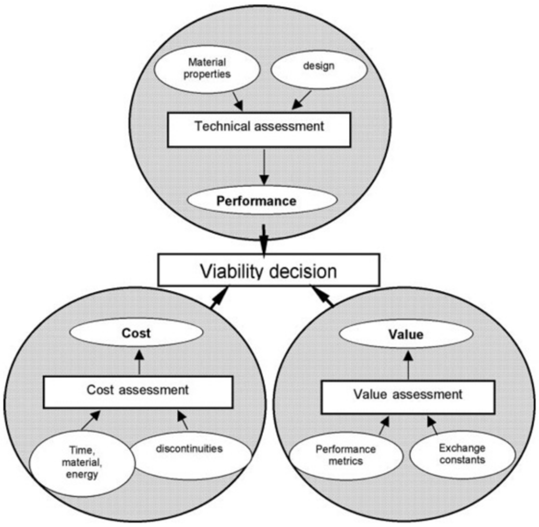


Figure 1
Three parts of a viability study.

2. *Technical Assessment*

The first step of a viability analysis is an application-specific performance comparison of the new material over the existing material by means of a technical model. This procedure should be applied to all promising applications. To construct a technical model, *performance metrics*, P_i, are identified and evaluated. A performance metric is a measure of the performance offered by the material in a particular application. In *minimum weight design* the performance metric is the mass: the lightest material that meets the specifications on stiffness, strength, etc., is the one with the greatest performance. In *design for energy mitigation*, in which it is desired that a protective packaging should crush, absorbing a specified energy, and occupy as little volume as possible, the performance metric is the volume. In *design to minimize heat loss*, the metric might be the heat flux per unit area of structure. In *design for the environment*, the metric is a measure of the environmental load associated with the manufacture, use, and disposal of the material.

Each performance metric defines an objective: an attribute of the structure that it is desired to maximize or minimize. Models are developed for the performance metrics following the methods detailed in Ashby (1999). The material that maximizes the chosen performance metric has the highest technical merit. But for this to be a viable choice, the performance must be balanced against the cost.

3. *Technical Cost Modeling*

The second step of a viability analysis is to establish the manufacturing cost differential between a component made with a novel process or material vs. existing processes and materials. If a predictive cost model has been created for manufacturing with a new material, then the model can be quickly used to compare the manufacturing cost of any part made with the new material with the known cost of the existing material part. *Technical cost modeling* (Field and de Neufville 1988, Clark *et al*. 1997) exploits the understanding of the way in which the control variables of the process influence production rate and product properties. It also uses information on the way the capital cost of equipment and tooling scale with output volume. These and other dependencies can be captured in theoretical and empirical formulas or look-up tables that are built into the cost model, giving greater resolution. In addition, informed sensitivity analysis and scenario building are enabled through the capturing of the linkages between the technical limitations of the process, intermediate variables such as cycle time, and cost line items.

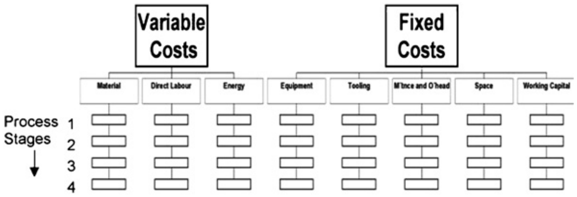


Figure 2
Schematic of a TCM.

A schematic of the structure of a technical cost model (TCM) is shown in Fig. 2. Each of the empty boxes represents the calculation of intermediate variables that capture the technical limitations of the process under varying input product specifications, process choices, and selected production volumes. The final output consists of both a range of unit cost numbers and the identification of cost drivers. As an example, we consider TCMs for two established processes for making *metal foams*. They are described briefly below.

3.1 *Liquid-state Foaming of Aluminum*

Consider a cost model for the production of panels of a SiC-stabilized aluminum-based metallic foam by the following four-step liquid-state process:

(i) melting of the premixed alloy;
(ii) holding, providing a reservoir;
(iii) foaming, using compressed gas and a bank of rotating blades; and
(iv) delivery of a continuous sheet of foam or of cut panels, via a moving belt.

The output of one step forms the input to the next, so the steps must match, dictating the size of the equipment or the number of parallel lines in each step. Data characterizing the dependence of desired material density and of production rate on the gas flow rate and the stirring rate have been incorporated in the model and links into the intermediate variables of limiting production rate, used hours per year, person-hours per year, fraction of capacity used, and number of parallel lines of equipment. Through these intermediate variables, the influences of how scale-up and varying product requirements affect the cost line items of equipment, direct labor, and power. Fixed costs are treated in the standard economics fashion, through the use of a cost-of-capital factor and by including the opportunity cost of capital associated with the working capital. A default range of product, process, and business variables can be overridden by the user, providing flexibility and the potential for extensive scenario and sensitivity analysis.

The outputs of the model (Figs. 3 and 4) show the way in which the cost of the manufactured component depends on production volume and also identify cost drivers. Significantly, the model indicates the production volume that would be necessary to reach the plateau level of cost at which, in the best case shown here, the material cost falls to roughly twice that of the input materials.

3.2 *Titanium Hydride Expansion of Aluminum via Powder Metallurgical Processing*

In comparison, consider a cost model for the production of panels of an aluminum-based metallic foam by either the batch or quasicontinuous processes for the powder metallurgical (PM) processing of aluminum foam. For the batch process, there are six steps:

(i) mixing of the aluminum powder with the powdered foaming agent ($<1\%$ TiH_2);

(ii) cold isostatic compression of the powder into an intermediate billet;

(iii) sintering of the billet (optional);

(iv) extrusion of the foamable precursor;

(v) placement of precursor in mold; and

(vi) batch foaming of precursor in mold.

For higher production volumes, automation of steps (ii), (iii), and (iv) and of steps (v) and (vi) have been proposed as follows:

(i) mixing of the aluminum powder with the powdered foaming agent ($<1\%$ TiH_2);

(ii) conversion of powder into precursor strip using the continuous CONFORM process; and

(iii) quasicontinuous foaming of precursor in the mold, meaning that the molds on a continuous belt are filled robotically, passed through the furnace, opened and, emptied.

The rate-limiting step for the low production volume process is the batch foaming step. The cycle time of this step is heat-transfer limited. The time required to heat the mold then increases with increased part dimensions and the time required to transfer heat into the precursor strips increases both with increasing part dimensions and with part curvature.

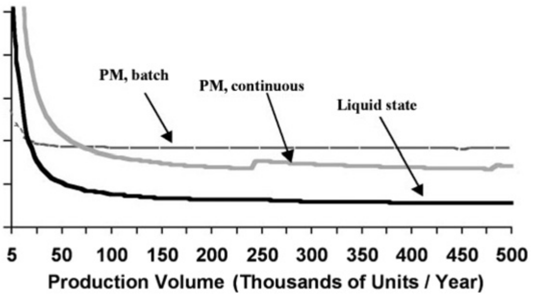


Figure 3
Manufacturing cost for varying production volumes: processing of aluminum foam by three methods.

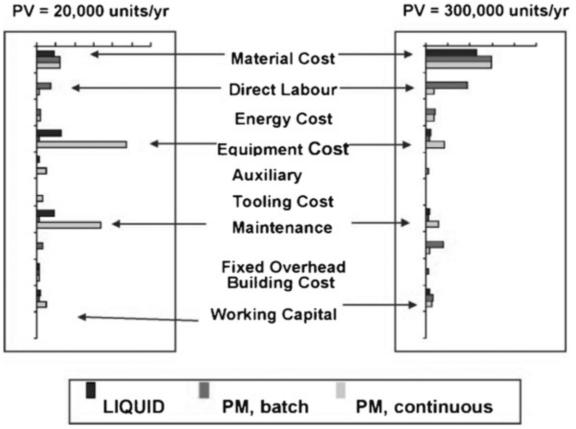


Figure 4
Line item cost for crossover annual production volume of 20 000 and 70 000 units: processing of aluminum foam by three methods.

An important cost driver in this process is the substantial amount of scrap generated in the foaming stage in order to ensure mold filling, shape definition, and to avoid weak zones along the interfaces of the precursor strips where they meet during foaming. The cost of processing the material lost during the foaming stage is reflected in the material cost line item of the batch foaming and continuous foaming stages. It is assumed that the scrap material can be sold as prompt scrap.

By running comparison scenarios with this model, it has been determined that the CONFORM process is not a viable alternative to the manual steps of billet pressing, billet heating, and extrusion of the precursor strips. However, a quasicontinuous foaming process to automate the rate-limiting batch process appears viable from this analysis. A comparison has been run for a simple flat panel of aluminum foam made by the three options of liquid-state processing, batch powder metallurgical processing, and the proposed quasicontinuous powder metallurgical process without the uneconomical CONFORM step. This simple flat panel comparison inherently favors the liquid-state process of Sect. 3.1, which is set up for the continuous production of flat panels. Figure 3 depicts the relative cost per unit of each of the different options over a range of annual production volumes. For annual production volumes of up to 20 000 parts, the batch powder metallurgical process is the most economical option. From examining Fig. 4, which shows a snapshot of the line item costs for each process at 20 000 and 300 000 parts per year, we can see that it is fixed costs, in particular equipment costs, that drive the low-volume cost of both continuous PM processing and liquid-state processing higher than that of batch PM processing. In this way, the TCM can be used to aid the decision about the suitability of converting to a continuous process.

By comparing the cheapest aluminum foam component (at the required production volume) with an existing material for any given application, the cost differential between a new and an existing material solution can be established. This provides input into the value assessment of the material.

4. Value Assessment

4.1 Multiobjective Optimization and Trade-off Surfaces

The viability of a foam in a given application depends on the balance between its performance and its cost. There may be several performance metrics that are important for the substitution decision, and the combination of these metrics and cost can be compared with those of the existing ones through the use of multiobjective optimization techniques.

When a design has two or more objectives, solutions rarely exist that optimize all objectives simultaneously. The objectives are normally noncommensurate, meaning that they are measured in different units, and in conflict, meaning that any improvement in one is at the loss of another. However, some solutions can be rejected quickly because they are dominated by other solutions: other solutions exist that have better values of both (or all) the performance metrics. The solutions that cannot be rejected in this manner lie on a line called the *nondominated* or *optimum* trade-off surface (Hansen and Druckstein 1982, Sawaragi and Nakayama 1985). The values of the performance metrics corresponding to the nondominated set of solutions are called the Pareto set (Fig. 5).

The trade-off surface identifies the subset of solutions that offer the best compromise between the objectives, but it does not distinguish between them. Three strategies are available to deal with this:

(i) The trade-off surface is established and studied, using intuition to select between nondominated solutions.

(ii) All but one of the objectives are reformulated as constraints by setting lower and upper limits for them, thereby allowing the solution that minimizes the remaining objective to be read off.

(iii) A composite objective function or *value function*, V, is formulated; the solution with the minimum value of V is the overall optimum. This method allows true multiobjective optimization, but requires more information than the other two. It is explored below.

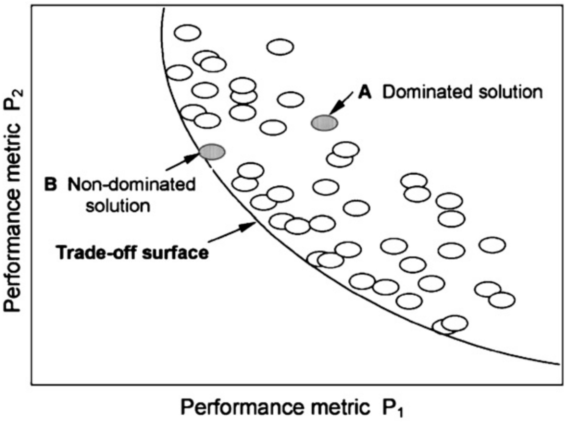


Figure 5
Dominated and nondominated solutions, and the optimum trade-off surface.

4.2 Value Functions

One strategy for optimizing over multiple objectives is through the use of a value function, V, defined below. For substitution to occur, the gain in value, ΔV, must be large enough to provide sufficient financial or strategic incentive to justify investment in the new technology, and also, if necessary, in the manufacturing facilities.

Value functions are used to compare and rank competing solutions to multiobjective optimization problems. The locally linear value function is defined as

$$V = \alpha_1 P_1 + \alpha_2 P_2 + \ \dots \ + \alpha_i P_i \ \dots \qquad (1)$$

where V is proportional to each performance metric P. The coefficients α are *exchange constants*: they relate the performance metrics P_1, P_2 ... to V, which is measured in units of currency ($, £, DM, etc). The exchange constants are defined by

$$\alpha_1 = \left(\frac{\partial V}{\partial P_1}\right)_{P_2,\ \ldots\ P_i} \tag{2}$$

$$\alpha_2 = \left(\frac{\partial V}{\partial P_2}\right)_{P_1,\ \ldots\ P_i} \tag{3}$$

that is, they measure the change in V for a unit change in a given performance metric, all others held constant. If the performance metric P_1 is mass m (to be minimized), α_1 is the change in value V associated with unit increase in m. If the performance metric P_2 is heat transfer q per unit area, α_2 is the change in value V associated with unit increase in q. The best solution is the one with the smallest value of V, which, with properly chosen values of α_1 and α_2, now correctly balances the conflicting objectives.

With given values of V and exchange constants α_i, Eqn. (1) defines a relationship between the performance metrics P_i. In two dimensions, this would plot as a family of parallel lines, overlaid on Fig. 5. The slope of the lines is fixed by the ratio of the exchange constants, α_1/α_2. The best solution is that at the point along a value line that is tangent to the trade-off surface, because this is the one with the smallest value of V. The method can be used to optimize the choice of material to fill a multifunctional role, provided that the exchange constants, α_i, for each performance metric are known.

(*a*) *Minimizing cost as an objective.* Frequently one of the objectives is that of minimizing cost, C, so that $P_1 = C$. Since we have chosen to measure value in units of currency, unit change in C gives unit change in V, with the result that

$$\alpha_1 = \left(\frac{\partial V}{\partial P_1}\right)_{P_2,\ \ldots\ P_i} = 1 \tag{4}$$

and Eqn. (1) becomes

$$V = C + \alpha_1 P_2 + \ \ldots\ + \alpha_i P_i \tag{5}$$

As a simple example, consider the substitution of a metal foam, M, for an existing (nonfoamed) material, M_0, based on cost C and one other performance metric, P. The value of P in the application is α. Substitution is potentially possible if the value V of M is less than that, V_0, of M_0. Thus substitution becomes a possibility when

$$V - V_0 = (C - C_0) + \alpha(P - P_0) \leqslant 0 \tag{6}$$

or

$$\Delta V = \Delta C + \alpha \Delta P \leqslant 0$$

from which

$$\frac{\Delta P}{\Delta C} \geqslant -\frac{1}{\alpha} \tag{7}$$

defining a set of potential applications for which M is a better choice than M_0.

To visualize this, think of a plot of the performance metric, P, against cost, C (Fig. 6). M_0 is centered at $\{P_0, C_0\}$; the potential substitute at $\{P, C\}$. The line through M_0 has the slope defined by Eqn. (7), using the equality sign. Any material that lies on this line, such as M_A, has the same value V as M_0; for this material, ΔV is zero. Materials above this line, such as M_B, despite having a lower (and thus better) value of P than M_0, have a higher value of V. Materials below the line, such as M_C, have a lower value of V (even though they cost more), a necessary condition for substitution. Remember, however, that while negative ΔV is a necessary condition for substitution, it may not be a sufficient one; sufficiency requires that the difference in ΔV be large enough to justify the investment in new technology.

As an example of the use of trade-off surfaces and value analysis, consider the performance–cost trade-off of aluminum foams vs. all other materials for a panel of specified stiffness. (This example is relevant in many automotive, aerospace, and infrastructure applications.) A cost metric is plotted along the vertical axis and a weight-adjusted stiffness performance metric along the horizontal axis. The Pareto set (shown by the trade-off line in Fig. 7) lies on a curved line owing to the logarithmic scale chosen to capture a wide range of materials. The three aluminum foams that lie on this trade-off surface are shown in black. The next step is to plot the value curves for different applications with different exchange constants onto this graph (as in Fig. 6), or evaluate them numerically. Table 1 illustrates the numerical method. Here

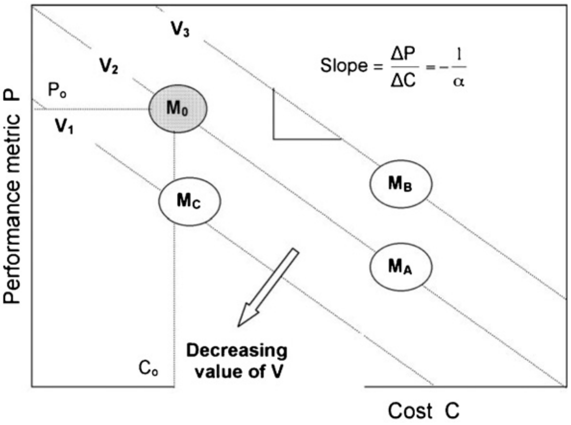


Figure 6
Trade-off between performance and cost.

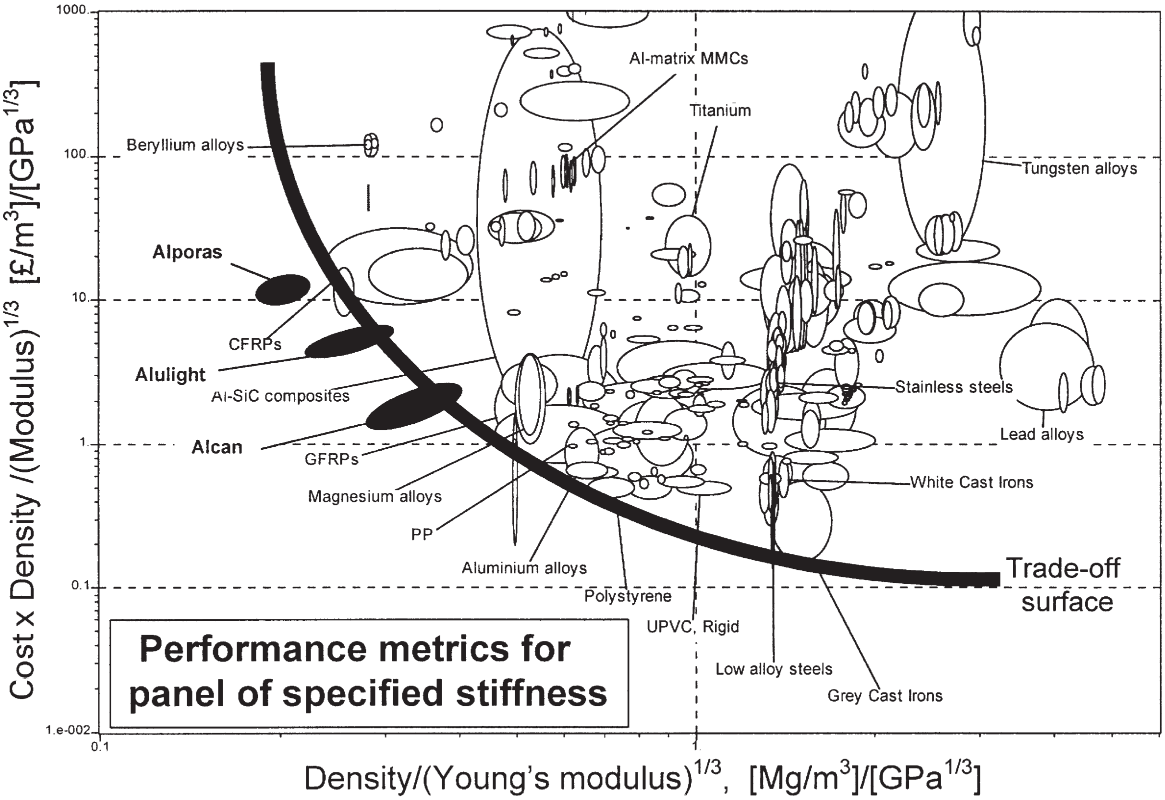


Figure 7
Performance–cost trade-off example of a light, stiff panel.

Table 1
The selection of panel materials: stiffness constraint.

Material	ρ (Mg m^{-3})	E (GPa)	C_m (£ kg^{-1})	P_1	P_2	V ($\alpha_1 =$ £0.5 kg^{-1})	V ($\alpha_1 =$ £500 kg^{-1})
Cast iron, nodular	7.30	175	0.25	1.31	0.33	0.99	655
Low-alloy steel (4340)	7.85	210	0.45	1.32	0.59	1.25	660
Al 6061-T6	2.85	70	0.95	0.69	0.66	1.01	345
Al-6061-20% SiC, PM	2.77	102	25	0.59	14.8	15.1	309
Ti-6-4 B265 grade 5	4.43	115	20	0.91	18.2	18.7	473
Beryllium SR-200	1.84	305	250	0.27	67.5	67.7	202
Alporas[a]	0.25	1.0	40	0.23	10.0	10.1	125
Alulight[a]	0.30	0.8	16	0.3	5.2	5.4	155
Alcan[a]	0.25	0.4	5.8	0.34	2.0	2.2	172

[a] All three types of metal foam are made in a range of densities, with a corresponding range of properties. These three examples are taken from the middle of the ranges. The costs are estimates only, broadly typical of current prices, but certain to change in the future. It is anticipated that large-scale production could lead to substantially lower costs.

P_1 is the weight-adjusted stiffness metric shown as the horizontal axis of Fig. 7, and P_2 is the cost metric shown as the vertical axis. For an exchange constant of $\alpha_1 =$ £0.5 kg^{-1}, cast iron has the lowest value of V and is the best choice; but for a larger exchange constant of $\alpha_1 =$ £500 kg^{-1}, the Alporas foam panel is a better choice than any other material.

4.3 The Exchange Constants, α

An exchange constant is a measure of the value of performance. Its magnitude and sign depend on the application. Thus the value of weight saving in a family car is small, though significant: in aerospace applications it is much larger. The use of heat transfer in a heat exchanger is positive; in applications requiring thermal insulation it is negative. The value of performance can be cost based, meaning that it measures an actual saving of cost, energy, materials, time, or information. But value can, sometimes, be perceived, meaning that the consumer, influenced by forces such as scarcity, advertising, fashion, aesthetics, convenience, etc., will pay more than the cost-based value of these metrics.

In many engineering applications the exchange constants can be derived approximately from technical models. Thus the value of weight saving in transport systems is derived from the fuel saving or the increased payload that this allows. The value of heat transfer can be derived from the price of the energy transmitted or conserved by unit change in this metric. Approximate exchange constants can sometimes be derived from historical pricing data; thus the value of weight saving in a bicycle can be approximated by plotting the price, P, of bikes against their mass, m, using the slope dP/dm as an estimate of α. Finally, exchange constants can be determined through interviewing techniques (Field and de Neufville 1988) that elicit the value to the consumer of a change in one performance metric, all others held constant.

5. Conclusions

Cost estimation is an important part of a material development strategy. A new material is viable in an application only if its value, so used, exceeds that of competing materials. Value is a function of both performance and cost and of one or more exchange constants that measure the use of performance. The value function that combines these allows optimized selection of material and process to meet a given design specification.

Acknowledgments

We wish to acknowledge the support of the Körber Foundation, the Cambridge Canadian Trust, and the UK EPSRC through the Engineering Design Centre at Cambridge, UK, and of Granta Design, Cambridge, UK, who developed the software illustrated in this article.

Bibliography

Ashby M F 1997 *ASM Handbook: Vol. 20. Materials Selection and Design*. ASM International, Materials Park, OH, pp. 281–90

Ashby M F 1999 *Materials Selection and Mechanical Design*. Butterworth Heinemann, Oxford, UK, 2nd edn

Clark J P, Roth R, Field F R 1997 Techno-economic issues in material science. In: *ASM Handbook: Vol. 20. Materials Selection and Design*. ASM International, Materials Park, OH

Field F R, de Neufville R 1988 Material selection—maximizing overall utility. *Metals Mater.* (June), 378–82

Hansen X, Druckstein Y 1982 *Multi-Objective Decision Analysis with Engineering and Business Applications*. Wiley, New York

de Neufville R 1990 *Applied Systems Analysis: Engineering Planning and Technology Management*. McGraw-Hill, New York

Sawargi Y, Nakayama H 1985 *Theory of Multi-Objective Optimisation*. Academic Press, New York

E. Maine and M. F. Ashby
Engineering Design Centre, Cambridge, UK

Mechanical Alloying

Mechanical alloying (MA) is a powder metallurgical process which was developed in the mid-1960s by the International Nickel Company (INCO) as a practical solid-state route for introducing and retaining fine refractory oxide particles in high-temperature alloys (Benjamin 1970, 1989). INCO established that it was possible to obtain the desired oxide distribution by using dry, high-energy ball milling of mixtures of raw material powders: mechanical alloying was born. It was quickly found that this solid-state powder processing route enabled intimately mixed, fine-scale composite powders to be produced from a wide range of elemental and master alloy starting powders. The principal aspects of powder processing and handling together with subsequent consolidation and thermomechanical processing which it was important to control in order to ensure reproducible end properties in high-temperature, creep-resistant nickel- (e.g., MA 754) and iron-based (e.g., MA 956) ODS alloys were also determined. Much of this processing knowledge has proved of generic relevance and is as applicable to the control of MA in the many other alloy systems that have subsequently been developed and explored as it was to the early MA ODS alloys.

An enormous amount of research has been performed on the MA technique and the materials systems that it has proved possible to synthesize using it. The broad categories for the latter include intermetallics, ceramics, composites (cermets and metal matrix composites (MMCs)), nanocrystalline materials, amorphous alloys, and self-propagating high-temperature synthesis (SHS) systems. The list is not exhaustive but at least illustrates the range of applicability and versatility of MA. Most of these applications of MA remain at the research stage.

1. Mechanisms and Processing

MA in high-energy ball milling equipment is accomplished by processing an initial powder charge usually comprising a mixture of elemental, ceramic (e.g., yttria for ODS alloys), and master alloy powders, all supplied in strictly controlled size ranges. Master alloy powders are used in order to reduce *in situ* oxidation of highly reactive species, such as aluminum or titanium alloy additions during processing. The milling medium normally used in commercial systems is a charge of hardened steel balls, typically 2 cm in diameter. The ball-to-powder weight ratio is chosen carefully for each mill and powder charge combination, but is typically around 10:1 for commercial systems. Given the enormous surface area, both of the initial powders and the fresh powder surfaces generated during MA processing, control of the milling atmosphere and its purity is essential to avoid undue alloy contamination. The principal protective atmospheres employed during commercial milling of MA powders are usually either argon or hydrogen and this protection generally extends both to pre- and post-MA powder handling. Both the purity of these gas atmospheres and the integrity of gas seals on the milling equipment are essential to control contamination, particularly when processing reactive species. For example, levels of oxide contamination in Ni_3Al can double with just a few hours of milling in impure argon. Occasionally, however, deliberate doping of the milling environment has been used to facilitate alloying additions during processing.

The central event during MA is the ball–powder–ball collision within the milling medium during processing. It is repetition of these high-energy collisions which leads eventually to MA of the powder charge. Intimate mixing and eventual MA of the powder charge occurs in a series of identifiable, more or less discrete stages during processing (e.g., Gilman and Benjamin 1983). For ductile–ductile or ductile–brittle combinations of starting powders, MA initially proceeds by the flattening and work hardening of ductile powders and fragmenting of brittle constituents, which is followed by extensive cold welding between powder particles, formation of lamellar structures, and coarsening of the powder particle size distribution. Brittle powder fragments are trapped at cold weld interfaces between the evolving lamellas of the ductile constituents and thus, while continuing to comminute, become dispersed. With continued milling a balance, which is dependent on processing parameters and the composition of the constituents, is established between further cold welding and powder particle fracture, leading to relatively stable powder particle sizes.

This balance between welding and fracture is accompanied both by further decreases in lamella spacings and by folding and mixing-in of lamella fragments to produce microstructures typical of MA (Fig. 1). For ODS alloys, powder constituents are milled to the stage where light microscopy examination reveals that lamella spacings have decreased to below the resolution limit ($\sim 1\,\mu m$). For typical levels of oxide addition (e.g., ~0.5wt.% yttria) this criterion ensures average dispersoid interparticle spacings of $<0.5\,\mu m$ (Fig. 2). In other systems, milling can progress until true alloying occurs. Surprisingly, MA can also be achieved between essentially brittle powder constituents. The mechanisms by which this occurs are less well understood than in systems incorporating at least one ductile powder component. Nevertheless, granular as opposed to interlamellar mixtures of brittle powder constituents do evolve, typically with smaller, harder fragments progressively incorporated to a very fine scale within the less hard constituents, e.g., alumina–nickel oxide. Moreover, MA of these brittle constituents can progress to true alloying, as has been demonstrated using lattice parameter measurements on Si–28 at.% Ge progressively milled from constituent powders (Davis and Koch 1987).

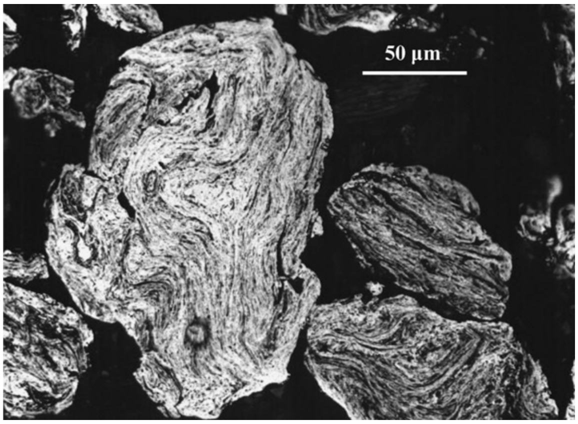


Figure 1
Polished and etched metallographic section of ODM 751 FeCrAl alloy powders in the fully MA condition, showing the folded lamellar structures typical of material processed by high-energy ball milling (courtesy of D.M. Jaeger).

Milling of very ductile metals such as aluminum and tin has to be carefully controlled to avoid complete agglomeration of the ductile phase rather than the balance between cold welding and fracture that leads to MA. This is normally achieved by adding precise amounts of organic compounds termed process control agents (PCAs) to the milling environment. Typically waxes or solvents, these compounds that interfere with cold welding progressively break down during milling to become incorporated within the final MA powders (e.g., in aluminum alloys) as fine-scale distributions of carbides or oxides. Similar restrictions to the proclivity for cold welding in ductile powders can be achieved without use of PCAs by milling at low temperatures, e.g., below −100 °C for aluminum.

The processing equipment used to effect MA by high-energy ball milling of powders originated in mining and conventional powder metallurgy industries.

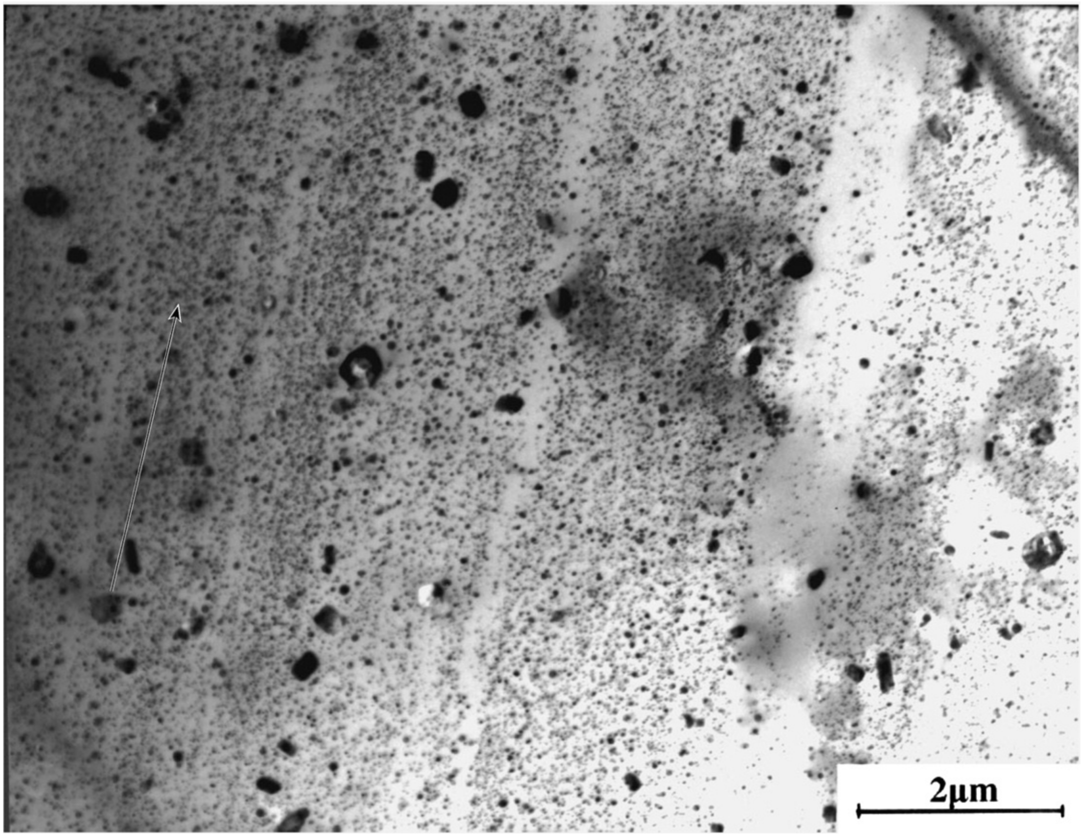


Figure 2
Transmission electron microscope image showing alignment of a fine-scale dispersion of oxide particles in extruded ODS alloy PM2000. The arrow shows the extrusion direction (courtesy of Y.L. Chen).

The range of high-energy ball milling equipment divides, approximately, into two categories: small, high-energy laboratory devices, and larger facilities capable of milling commercial quantities of powder. The former category includes SPEX shaker mills and planetary ball mills. Both devices are capable of rapidly effecting MA, but in quantities of powders up to no more than a few tens of grams. SPEX mills vibrate at up to 1200 rpm in three orthogonal directions to achieve ball velocities approaching $\sim 5\,\mathrm{ms}^{-1}$. Planetary mills incorporate a rotating base plate upon which are mounted counter-rotating, smaller-radius vials containing the ball/powder charge. The kinetic energy imparted to the ball charge in the planetary mill depends on the base plate and vial radii and angular velocities. Attritor or Szigvari ball mills, depending on their size, can be used either for laboratory or commercial ball milling applications and incorporate a rotating vertical shaft with attached horizontal impellors which stirs a container housing the ball and powder charge. These devices can process batches of up to several kilograms or more of powder through the significant differential movement the impellors generate between the ball and powder charge. Balls can either cascade or tumble when leaving the mill wall during attritor processing, depending on the ball charge and impellor velocity.

The largest commercial devices applied to MA are horizontal ball mills. When these devices exceed several meters in diameter they impart sufficient kinetic energy through ball impacts to effect MA and can process over 1000 kg of powder per batch in larger units. Balls either cascade or tumble during processing in these mills depending on rotational speed (see Fig. 3). The time taken to achieve MA scales approximately inversely to the size of the milling equipment used. Hence, milling which might take minutes to accomplish in a SPEX mill could take hours in an attritor or days in a horizontal ball mill. All of these processing routes, however, have very low energy conversion efficiency, in that only a small fraction of the milling energy expended effects microstructural change contributing to the MA process.

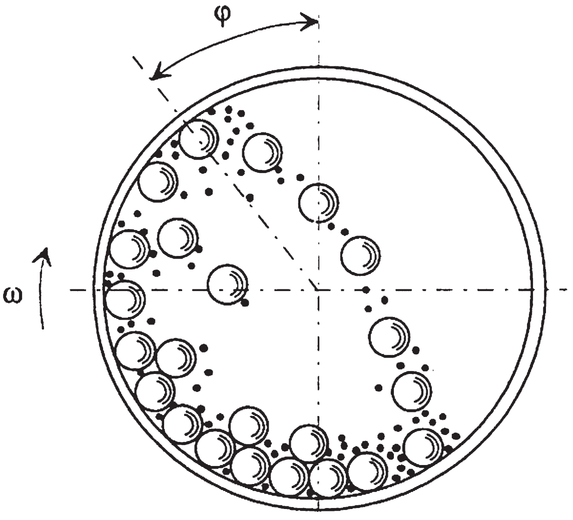


Figure 3
Configuration of a horizontal ball mill, showing the release of the powder and ball charge (at angular position φ) from the inner wall of the mill rotating with angular velocity ω (after Lu *et al.* 1995).

It is worth noting that during MA, powder particles also coat (condition) the ball milling medium, which means that, to avoid cross-contamination of commercial alloys, the repeat use of ball charges is restricted to compositionally similar batches of raw materials.

2. Modeling

The approaches taken to modeling the mechanical and dynamical aspects of MA fall into two broad categories: local and global modeling (Courtney 1994, Soni 2000). Local modeling considers the likely fundamental nature of single-collision events, the energy transfer to trapped powder particles, and the way particle welding and fracture events may lead to temporal evolution of key powder parameters such as hardness, particle shape, and size distribution. In order to effect the computations necessary to follow iteratively the evolution of powder particle changes during MA, local modeling requires input of mill-specific parameters such as ball velocity and collision frequencies determined from global modeling approaches. Analytical approaches to local modeling have been based on assumptions of Hertzian collisions between balls leading to transfer of energy to trapped powder particles that deform in a manner that can be estimated from constitutive equations for plastic deformation. These models predict strain rates in the deforming powders of $\sim 10^3$–$10^4\,\mathrm{s}^{-1}$ with accumulated strain per collision ranging from a few tenths of 1% (attritor mill) to 1% (SPEX mill).

Calculations of temperature change in MA powders during milling, assuming either adiabatic heating or temperature rise owing to Hertzian pressure, predict modest rises (several hundred kelvin in a SPEX mill) in line with observation. These local models of MA have been extended to consider particle cold welding and particle fracture. The latter events are modeled by assuming a mechanism similar to that of a forging fracture, where this can be taken to originate at the collision impact point and spread radially across the flattened powder under the appropriate conditions of crack initiation and propagation (J-integral approach). Using mill parameters estimated from global modeling, local analytical models have been incorporated in computational programs and run to predict the time-based evolution of powder size, hardness, and microstructural refinement occurring during MA.

Comparisons with experimental data obtained from elemental chromium and iron powders processed in a SPEX mill, reveal that the model is capable of predicting the correct shape of curve for the evolution of the average particle size with milling time but underestimates the time to achieve peak average particle size and overestimates its value. Other predictions are for lower hardness and less microstructural refinement than are observed experimentally. Similar predictions for aluminum milled in a SPEX unit again underestimate particle hardness but achieve closer agreement with evolving particle sizes. It is generally accepted that uncertainties in the values of key parameters used both in these and in more phenomenological, local models of MA mean they must be considered qualitative in nature. However, they enable reasonable predictions of trends when either powders or processes are changed systematically.

The requirement for global modeling of MA reflects the inherent complexity of the high-energy ball milling process, added to which are device-specific characteristics such as positional milling and powder inhomogeneities. For predictive purposes global models of MA seek to define ball impact frequencies, angles, and velocities together with estimates of the effective fraction (that leading to energy transfer likely to be of significance in terms of MA) of milling impact events. Much of this work has been underpinned by experimental validation of patterns of local milling behavior and specific effects of mill design using techniques such as high-speed photography. These models have revealed, for example, that $<1\%$ of ball impacts in a SPEX mill involve the near head-on collisions most effective for ball–powder energy transfer; and that increasing the ball charge increases the impact frequency, while simultaneously decreasing both ball collision velocity (lower mean free path) and the final hardness of milled powders. Similar results have been obtained for analysis of planetary ball mill action, although here it has been established that two different milling regimes (energies) exist, depending on the dimensions and respective angular velocities of support plate and vial. These are ball rolling on the vial inner surface, and ball escape and impact on the vial opposite wall. By independently varying plate and vial angular velocity and thus ball energy, Gaffet (1991) demonstrated that it was possible to transform a powder charge with an average composition of $Ni_{10}Zr_7$ to end states ranging between fully crystalline and fully amorphous, thus underlining the link between process and outcome, or MA and microstructure.

One disadvantage of local/global modeling of the dynamics and mechanics of MA is that these models are not linked to fundamental transformation mechanisms, particularly important for understanding the formation of metastable phases during MA. However, atomistic models of metastable phase formation based, for example, on molecular dynamics are unconnected to the deformation parameters of MA (Khina and Froes 1996). Both phenomenological thermodynamic (e.g., based on heat of mixing) and kinetic (based on multiple layer diffusion couples) models have been used to interpret this metastable phase formation, but further developments of these models are needed to link plastic strain and defect levels directly to the kinetics as well as nucleation and growth mechanisms of metastable phases formed during MA.

3. *Applications*

Despite the fact that MA ODS alloys have been used successfully since high-strength, corrosion-resistant, nickel-based INCO alloy MA 754 bar was used by GEC in the nozzle guide vane assembly of the F404-400 military fighter aircraft engine in 1977, the high cost of these alloys has determined that they remain niche materials. Essentially, the only MA materials that have successfully penetrated commercial markets so far are ODS alloys, principally the nickel- and iron-based variants. Total worldwide production of these materials remains at a level of around several hundred tons per year. Several aluminum-based ODS alloys were developed by INCO from the late 1970s through the 1980s culminating in the commissioning of a production facility in 1988 that was closed in 1992 when INCO withdrew MA ODS aluminum alloys from the market. Potential aerospace end users of INCO high-strength, as-forged, lightweight, stress corrosion-resistant Al–Mg–Li ODS alloy, designated Al-905XL, had to wait seven years for a viable replacement to emerge (designated alloy 5091), produced by Aerospace Metal Composites Ltd.

Nickel-based ODS alloys range from the solid solution-strengthened ODS Ni–20 wt.% Cr alloy MA 754 to alloy MA 758, which is essentially similar but contains 30 wt.% chromium (for corrosion resistance in components used for manufacturing fiberglass from the melt), to the γ' and solid solution (tungsten, molybdenum) strengthened, creep, oxidation, and sulfidation (chromium, aluminum, titanium, tantalum, and tungsten) resistant ODS alloys MA 6000 and MA 760. These higher strength nickel-based ODS alloy variants have found applications as small, highly stressed first- and second-stage turbine blades (MA 6000) and for industrial gas turbine vanes and blades (MA 760). These latter alloys are supplied for service in a very coarse-grained, high grain aspect ratio condition generated by dynamic zone annealing that imparts excellent high-temperature creep resistance. The nickel-based ODS alloys are more difficult to form than their iron-based ODS alloy counterparts and are available in a more limited range of mill forms (e.g., bar, wire-rod, plate, sheet, and precision forgings).

Iron-based ODS alloys (MA 956, PM 2000, and the higher chromium, aluminum variant MA 956HT) which are, essentially, dispersion-strengthened, highly oxidation-resistant variants of FeCrAl alloys, have found a greater variety of specialist applications in a broader range of market sectors than their nickel-based ODS alloy counterparts. As with nickel-based ODS alloys, MA 956 and PM 2000 achieve maximum creep and oxidation resistance when used in service in the high aspect ratio, coarse-grained secondary recrystallized, preoxidized release condition, achieved for iron-based ODS alloys without resort to zone annealing. In contrast, the iron-based MA ODS alloy MA 957, which was developed for use as a low void swelling nuclear fuel cladding tube in fast breeder reactors and is intended for service in liquid sodium at $\sim 700\,^{\circ}C$, is supplied with a specified $\sim 30\,\mu m$ grain size. The excellent fabricability of iron-based ODS alloys ensures they are amenable to a wide range of metal processing techniques.

Other niche applications, for MA aluminum alloys, include race pistons for motorbikes forged from Al–12 wt.% Si, DISPAL-1, produced by reaction-milling techniques, and thin-walled tube for road racing bicycle frames extruded from hot isostatically pressed (HIP) billets of 2124-T4-17.8 wt.% SiC_p alloy, manufactured by MA. For both these alloys as well as for the low-density 5091 aerospace alloy, service applications are at ambient, hence the material is supplied in a very fine-grained condition (e.g., $< 1\,\mu m$ grain size for 5091 alloy).

Elsewhere, essentially small amounts of special composition MA materials have been tested in niche applications, such as the aluminum alloy DISPAL-2 used in electrodes for NaS batteries or for low-distortion aerospace interferometer parts, and the magnesium-based "supercorroding" alloy exhibiting controlled corrosion rates for application as bolting to release deep-sea equipment or as controlled heating/hydrogen sources. In the area of nonstructural applications, a number of specialist coatings have been produced using MA powders. These include Ti–Al powders for RF plasma spraying, Ti–Al–SiC powders for wear-resistant coatings, and sieved, as-MA powders of the standard commercial ODS alloys such as MA 754 and PM 2000 for plasma spraying. A promising emerging MA ODS alloy is the intermetallic composition, yttria-strengthened alloy Fe_3Al, intended for application in components for use in fossil fuel power plants operating above 1100 °C, where the high aluminum content in the alloy is intended to confer enhanced oxidation resistance in aggressive environments.

Future commercial opportunities for MA alloys include the possible use of amorphous Al–Cu–Mg alloys as aerospace structural materials, intermetallics such as Ti_3Al, TiAl, and Cr_3Si for either airframe or high-temperature engine applications, and MA magnesium-containing combinations of Fe/Ni/Co/Ti as potential hydrogen storage materials (Lu and Lai 1998, Soni 2000).

See also: Mechanical Milling#

Bibliography

Benjamin J S 1970 Dispersion strengthened superalloys by mechanical alloying. *Metall. Trans.* **1**, 2943–51

Benjamin J S 1989 Mechanical alloying—a perspective. In: Arzt E, Schultz L (eds.) *New Materials by Mechanical Alloying Techniques: Proc. DGM Conf.* Deutsche Gesellschaft für Metallkunde, Germany, pp. 3–18

Courtney T H 1994 Modeling of mechanical milling and mechanical alloying. *Rev. Particulate Mater.* **2**, 63–113

Davis R M, Koch C C 1987 Mechanical alloying of brittle components: silicon and germanium. *Scr. Metall.* **21** (3), 305–10

Gaffet E 1991 Planetary ball milling. An experimental parameter phase diagram. *Mater. Sci. Eng. A* **132** (1–2), 181–93

Gilman P S, Benjamin J S 1983 Mechanical alloying. *Annu. Rev. Mater. Sci.* **13**, 279–300

Khina B B, Froes F H 1996 Modeling mechanical alloying: advances and challenges. *J. Mater.* **48** (7), 36–8

Lu L, Lai M O 1998 *Mechanical Alloying*. Kluwer, Boston

Lu L, Lai M O, Zhang S 1995 Modeling of the mechanical alloying process. *J. Mater. Process. Tech.* **52**, 539–46

Soni P R 2000 *Mechanical Alloying—Fundamentals and Application.* Cambridge International, Great Abington, Cambridge, UK

Suryanarayana C 1999 Mechanical alloying. In: Suryanarayana C (ed.) *Non-equilibrium Processing of Materials.* Pergamon, Oxford, UK, pp. 49–85

Suryanarayana C 2001 Mechanical alloying. *Prog. Mater. Sci.* **46**, 1–184

A. R. Jones
University of Liverpool, UK

Mechanical Testing Methods of Fibers and Composites

Commonly employed techniques to determine the mechanical properties of fibers and fiber-reinforced metal- and polymer-matrix composites (MMCs and PMCs, respectively) are briefly described in this article, some of which are also applicable to ceramic-matrix composites (CMCs). Wide coverage of these test methods is given in ASTM standards (ASTM Standards 1998) and in sections of this encyclopedia on composite materials (see *Composite Materials: Overview#*).

1. Fibers

The properties of interest at low temperatures are the elastic modulus and the fracture strength whereas for high temperature applications and long-term property modeling, creep properties such as the creep–rupture strength, time- and strain-to-failure are necessary. Most fiber properties are probabilistic in nature, hence, many fibers have to be tested and the data generated should be statistically analyzed. It is recommended that the fiber properties be measured on a material that represents composite behavior. For example, fibers extracted from as-processed or fatigued composite specimens should be tested to obtain properties after processing or fatigue testing, respectively.

The standard method for determining room temperature properties is the single-filament tensile test (ASTM D3379). In this method, a random selection of fibers is made from the material to be tested. The fibers are centerline-mounted on special slotted tabs (Fig. 1). The tabs are clamped so that the test specimen is aligned axially in the grips of a test machine and then loaded to failure at a constant displacement rate. It is important to measure the cross-sectional area carefully. The axial strain is measured using a noncontact- type extensometer.

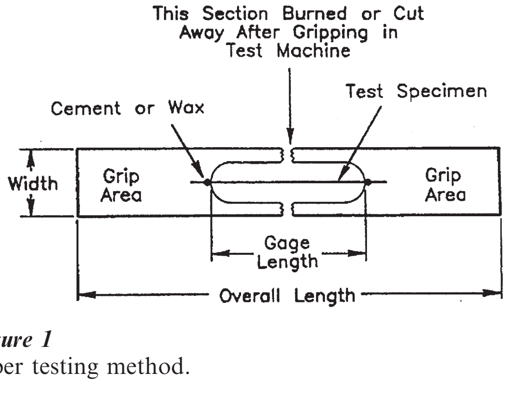


Figure 1
Fiber testing method.

For carbon and graphite yarns, strands, rovings, and tows, the resin impregnated tow tensile test (ASTM D4018) is recommended. The purpose of the impregnating resin is to provide the tow with sufficient mechanical strength to produce a rigid test specimen capable of sustaining uniform loading of the individual filaments. Care should be taken to align the fibers prior to impregnation. A resin compatible with the fibers and in minimum required quantity to produce the tensile specimen should be used. The failure strain of the resin should be significantly higher than that of the fibers. End tabs (cast or mounted) may be necessary in some instances for alignment purposes and to prevent slippage at high loads.

Creep testing of fibers is carried out using dead load tests at constant temperatures (either in air or in a suitable environment). A length of fiber is gripped vertically in cold grips and a resistance furnace is used to maintain constant temperature over a specified gauge length (typically 25 mm) of the fiber. Elongation and fracture strain are measured using any of the noncontacting displacement devices. Bend stress relaxation is used for measuring the creep behavior, particularly for small diameter fibers. This method consists of tying the fiber into a loop and exposing it to the appropriate temperature for a specific time. Afterwards, the diameter of the loop is measured before breaking it. The residual loop radius is used to estimate the creep behavior.

2. *Composites*

2.1 Tension

Straight-sided dog-bone type loaded specimens with rectangular cross-section are commonly used for measuring the in-plane tensile properties of unidirectional, as well as cross-ply, composite laminates (ASTM D3039 for PMCs and D3552 for MMCs). Specimens can be end-tabbed if necessary to distribute the load from the grips (either wedge or hydraulic) into the specimen with a minimum stress concentration. Test specimens with round or tubular cross-sections can also be used. An important factor for obtaining highly accurate test results is minimization of the specimen misalignment with the load train and the resultant bending stresses. Likewise, good alignment between the desired fiber orientation and the specimen axis is a must.

The following alternative methods are also available.

(i) The sandwich beam test (ASTM C393) for evaluating the flexural properties of flat sandwich constructions in which the test laminate is bonded to either side of a suitable honeycomb core and the sandwich is subsequently tested in a four-point flexure configuration. To minimize thermal expansion mismatch problems, the thickness of the laminate on the compression side should generally be twice that of the tensile side. This test is also useful for tensile testing, particularly for the 90° properties of unidirectional composites.

(ii) Split-disc ring tensile test for apparent tensile strength of ring or tubular composites (ASTM D2290) in which a hoop-wound narrow ring is loaded using a split-disc loading fixture that applies a hoop stress to the test ring.

The out-of-plane tensile properties of composites can be evaluated by direct out-of-plane loading using a laminated specimen bonded between two fixture blocks (square or circular cross-sections) that are pulled in tension. Alternatively, curved composite beams, which take advantage of the out-of-plane tensile loading induced in the elbow of a curved laminate beam subjected to an opening moment, can be used for indirect measurements.

2.2 Compression

Compressive testing of composites is relatively more difficult than tensile testing. The mode of failure involves some type of buckling, ranging from classical column buckling of the entire specimen to local micro buckling of the individual fibers. The test fixture should be appropriate to avoid buckling. Many axial compression test methods in current use are some variation of the Celanese compression test (ASTM D3410). In this method, a thin, straight-sided specimen with a small gauge length (~ 13 mm) is loaded in the tabbed region using inverted cone wedge grips. The IITRI method (ASTM D3410), which is a minor modification, incorporates flat wedge surfaces rather than conical surfaces for greater stability. An untabbed, thick, straight-sided composite specimen can also be used, end-loaded in compression and side-supported to control lateral buckling (ASTM D695). The end-loading might induce deformation and minor crushing at the ends, however these are not of much concern if strain gauge or extensometers are used to monitor strains rather than cross-head displacement. This method has the advantage of requiring a simpler test specimen and test fixture than the Celanese configuration. The compression specimen geometry and method of load introduction should be relevant to the actual application for which the data is being generated.

2.3 Shear

The in-plane shear response of the composites can be obtained by tensile testing of a $\pm 45°$ laminate (ASTM D3518). This test is simple and requires little fixturing. The Iosipescu shear test (ASTM D5379) utilizes a V-notched beam wherein a material coupon in the form of a rectangular flat strip with symmetrical centrally located V-notches is loaded using a special test fixture with the notches located along the line of action of loading (Fig. 2). Two strain gauges,

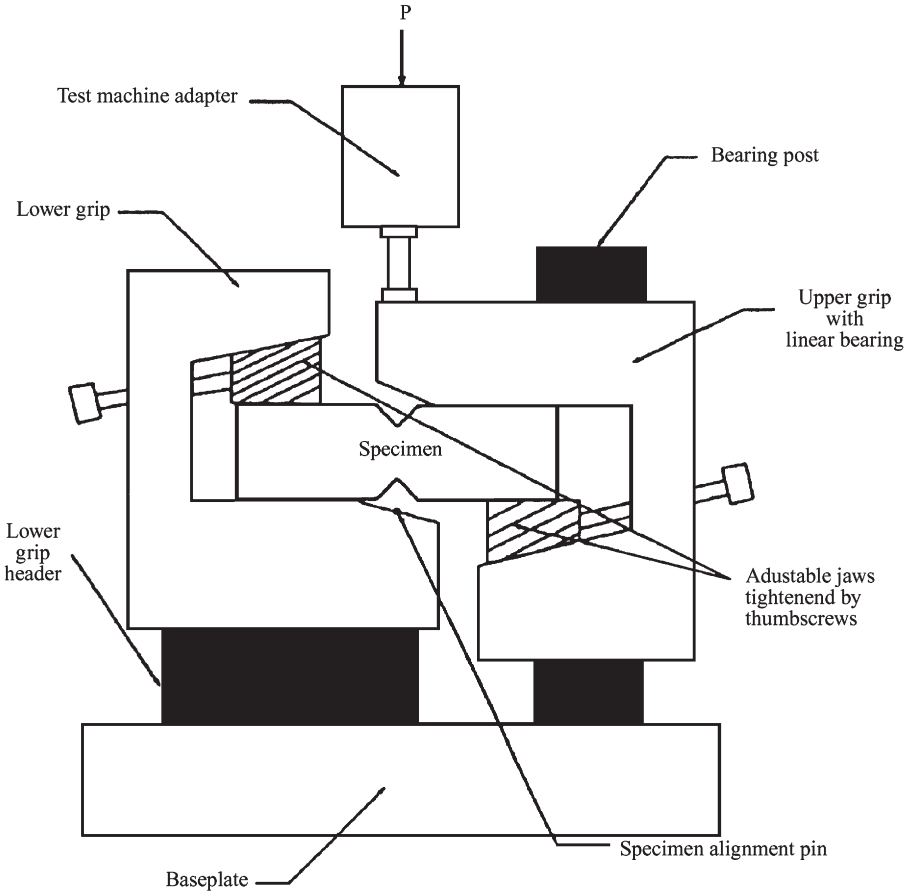


Figure 2
The Iosipescu shear test.

oriented at ±45° to the loading axis and located in the middle of the specimens and along the loading axis, are used for measuring the shear response. By changing the material coordinate system with respect to the loading axis, both in-plane and out-of-plane shear properties can be evaluated using this technique. Additional methods such as the rail shear test (ASTM D4255-83), the ten-degree off-axis shear test, and the tube-torsion test (ASTM D5448) are also available. The out-of-plane shear properties can be evaluated by using the short-beam strength test (ASTM D2344) or the Iosipescu test.

2.4 *Fracture Toughness*

The double cantilever beam (DCB) test (ASTM D5528) is the most widely used test method for measuring Mode I fracture toughness of unidirectional composites. The specimens are ~125 mm long, 20–25 mm wide, and 3–5 mm thick. One side of the specimen is notched (~25 mm deep) and the cracked faces pulled apart with the aid of either piano hinges or loading blocks attached to the specimen on the cracked end. Load–displacement data are recorded and used for computation of the energy release rates. For Mode II toughness, the end notched flexure test, which uses a specimen similar to that of the DCB test but loaded in three-point bending configuration, can be used.

2.5 *Fatigue*

Fatigue in composites is application dependent, hence the laminates should represent the application and testing should account for the loading spectra and environment experienced during service. Unnotched fatigue tests on specimens similar to those used for tension tests can be conducted, either under load or strain control and either triangular or sinusoidal waveforms (ASTM standard D3479 for PMC and E466 for MMC). High loading frequency (>10 Hz) may lead to frictional heating at the fiber/matrix interface due to relative sliding. Fatigue crack growth tests are conducted using either center-notched or single edge-notched specimens, preferably with fixed far-field stress amplitude. Compact-tension (CT) geometry can also be used if the composite laminates have sufficient thickness. The data are typically presented as crack extension versus the number of fatigue cycles. By use of suitable fiber bridging models, the data can also be converted into crack growth rates against the effective stress intensity factor range.

3. Interfacial Properties

Fiber/matrix interface properties such as the debonding strength and sliding resistance play a predominant role in determining the properties of the composites. Fiber push-out testing is a commonly employed technique for measuring these properties. In this test, an indenter (punch) is used to apply axial compressive loading to a selected fiber in order to debond the fiber and force it to slide relative to the matrix. The specimen, a thin slice of the composite (typically 0.6–1.3 mm thick) with the fibers oriented axially within ±1°, is mounted on a support plate that has grooves to facilitate the fiber once it has been pushed out from the opposite face of the sample. A detector can be used to monitor acoustic emissions associated with the debonding events. Composites that have high interfacial strengths are not amenable to this technique. For composites that exhibit single or multiple matrix cracking, alternative techniques include single fiber pull-out and hysteresis measurements.

Bibliography

Annual Book of ASTM Standards 1998 Vol. 15.03, American Society for Testing and Materials, Philadelphia, PA

U. Ramamurty
Indian Institute of Science, Bangalore, India

Membrane-based Polymer–Ceramic Nanocomposites

Membrane-based polymer–ceramic nanocomposites are a category of organic–inorganic hybrid materials. As their name indicates, these materials have a polymer membrane as the organic phase and a ceramic as the inorganic phase. To be clearer, the polymer membrane serves as a matrix, and ceramic materials are introduced into it as described below.

Significant progress has been made in research on membrane-based polymer–ceramic nanocomposites. Preparation methods have been developed for the fabrication of nanocomposites (Shao *et al.* 1995, Wen and Mark 1995a, 1995b, Apichatachutapan *et al.* 1996, Gummaraju *et al.* 1996, Deng *et al.* 1998a, 1998b). Also, the properties of nanocomposites have been investigated by different characterization techniques (Wen and Mark 1995a, 1995b, Deng *et al.* 1998a, 1998b, Young and Mauritz 1998, Start and Mauritz 1998, Siuzdak and Mauritz 1999). Applications for nanocomposites in a variety of areas have been explored as well. Presented here are the synthetic approaches, properties, and possible applications of membrane-based polymerceramic nanocomposites. The outlook for future research is also discussed.

1. Synthetic Methods

Great synthetic challenges arise in preparing organic–inorganic hybrid nanocomposites. Owing to the strong tendency of nanoparticles to agglomerate, organic–inorganic hybrid nanocomposites are not usually obtained by mechanical methods, and chemical routes have been developed instead. Techniques used to introduce, assemble, or organize the individual nanoparticle entities into polymers have included dispersion into polymers, direct synthesis in micellar polymers, and absorption into polymer layers (Hadjipanayis and Siegel 1993, Kunz *et al.* 1993, Roecher and Moller 1995, Spatz *et al.* 1996, Lerner and Oriakhi 1997). These techniques typically involve solution phase processes, with strategies of either (i) the nucleation or formation of an inorganic structure in a solution containing or embedded in the desired polymer or (ii) monomer insertion and subsequent *in situ* polymerization within an inorganic host framework. For membrane-based polymer–ceramic nanocomposites the first strategy, specifically, an *in situ* sol–gel approach (Fig. 1), is currently the most widely utilized.

To implement this sol–gel approach the following steps are required:

the polymer film is preprocessed to remove low molecular weight fractions or other unwanted materials;

the polymer film is swollen in a solvent (alcohol, toluene, mixed solvents, etc.) to make it able to or easier to accommodate precursors;

the polymer film is placed into a precursor solution to let the precursor penetrate, and the sol–gel reaction is initiated and implemented;

the nanocomposite film is postprocessed (dried, placed under vacuum, etc.).

1.1 Materials for the Ceramic Phase

Many ceramic materials are good candidates for membrane-based polymer–ceramic nanocomposites. Owing to the difficulties encountered in some synthetic

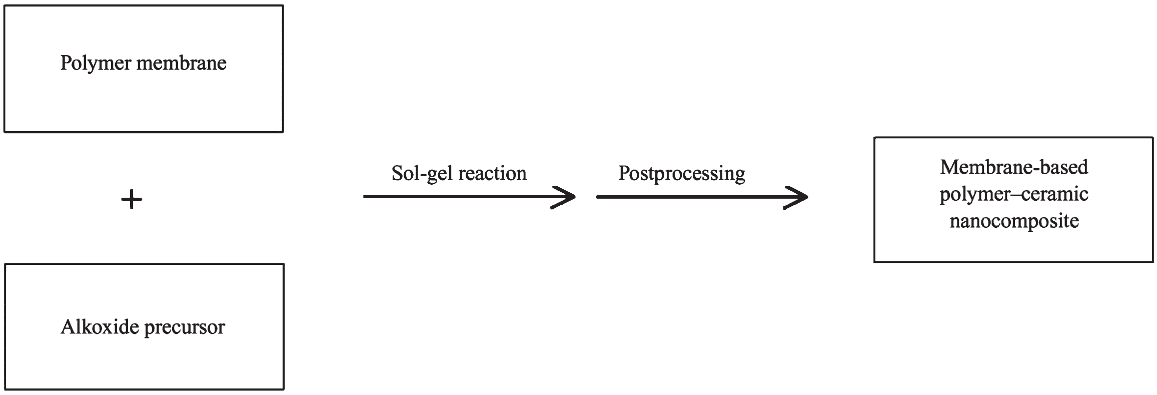


Figure 1
The currently employed synthetic approach for the preparation of membrane-based polymer–ceramic nanocomposites.

Table 1
Some membrane-based polymer–ceramic nanocomposite systems

Polymer membrane	Ceramic phase	Synthetic method	References
Poly(ethylene-co-methacrylic acid) (Surlyn)	Silicon oxide	Sol–gelreaction	Start and Mauritz (1998)
Perfluorosulfonate (Nafion)	Titanium oxide/silicon oxide, Aluminum oxide/silicon oxide, Organically modified silicate (ORMOSIL)	Sol–gelreaction	Shao *et al.* (1995), Young and Mauritz (1998) Deng *et al.* (1998a)
Poly(styrene-*b*-isobutylene-*b*-styrene)	Silicon oxide	Sol–gel reaction	Mountz *et al.* (1998)
Silicone	Silica, alumina, titania, zirconia, and mixed oxides	Sol–gel reaction	Wen and Mark (1995)

approaches, currently the most widely researched materials are oxides, organically modified oxides, and mixtures of oxides. Examples are silicon oxide, aluminum oxide, titanium oxide, and zirconium oxide. Consequently, the materials used in precursor are generally an alkoxide or a mixture of alkoxides.

1.2 Materials for the Polymer Phase

The polymer phase serves as a matrix to host the ceramic particles. In the sol–gel synthetic method the polymer membrane needs to be swollen by certain solvents, so it must be able to tolerate the solvents and precursor solutions. It will be ideal if the polymer is cross-linked. Also, it is desirable that the polymer phase has good compatibility with the ceramic phase to facilitate good ceramic uptake. Examples of polymer membranes are silicones, Nafion, biopolymers, and many others, some of which are listed in Table 1.

2. Characterization Methods

Many traditional characterization methods can be used to investigate the general properties of membrane-based polymer–ceramic nanocomposites, and some special analytical techniques are available for specific properties. These include the stress–strain test and dynamic mechanical analysis for mechanical properties, thermogravimetric analysis, differential scanning calorimetry for thermal properties, infrared spectroscopy, and nuclear magnetic resonance for chemical structure and bonding, transmission electron microscopy for the imaging of inorganic particles, and environmental scanning electron microscopy for the exploring of element concentration profiles across the film cross-section.

3. General Properties

3.1 Percentage Uptake of Ceramics in a Composite

The percentage uptake of ceramics in a membrane-based nanocomposite depends on many factors: the thickness of the polymer membrane, the time allowed for swelling in the solvent, the length of time when soaking in the precursor solution, the compatibility between the polymer film and the precursor and ceramic materials, and so on. Research shows that most of the uptake of ceramic materials in nanocomposites occurs during a short initial period of time, longer times only increasing the uptake slightly.

3.2 Distribution Profiles of Ceramics across the Polymer Membrane Cross-section

The sol–gel process of generating ceramic nanoparticles within a polymer membrane requires the precursor solution to penetrate into the membrane. Due to the thickness of the membrane and the possible molecular incompatibility between the membrane and the precursor solution, even when a sufficiently long period of time has been allowed for the swelling process, the ceramic nanoparticles will not be evenly distributed across the cross-section of the polymer membrane. Research shows that the concentration of ceramic materials is the highest at and near the surface and drops sharply over the thickness. Figure 2 illustrates such a typical profile.

3.3 Mechanical Properties

The loading of ceramic materials into polymer films has a marked impact on the mechanical properties of the resultant membrane-based polymer–ceramic nanocomposites, in comparison with the properties of conventional composites (Mark *et al.* 1995). Generally speaking, compared with its corresponding original polymer membrane, the polymer–ceramic composite has improved mechanical strength, but lower elongation to break. Yet nanocomposites have their own advantages. Owing to their nanometer-scale size, the ceramic particles in these materials have larger surface areas, and therefore better binding or adhering interactions with a polymeric matrix than those in composites prepared from blending or from micrometer-sized particles. These interactions help reduce the mobility of the polymer chains and improve toughness and other mechanical properties.

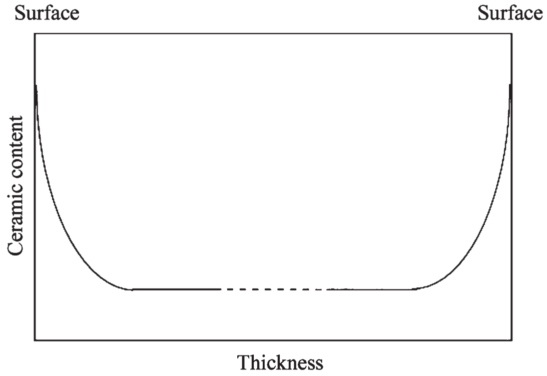


Figure 2
Typical distribution profile of ceramics across the polymer membrane cross-section.

3.4 Thermal Properties

The thermal properties of membrane-based polymer–ceramic nanocomposites benefit from the *in situ* sol–gel synthetic process. On one hand, nanocomposites have good thermal stability because the ceramic phase can be generated at low or even room temperature rather than at the elevated temperatures required in the production of conventional ceramic materials (Brinker and Scherer 1990, Hench and West 1990). On the other hand, nanocomposites have higher glass transition temperatures (T_g) than the corresponding polymeric membranes.

3.5 Other Functional Properties

The polymer phase in membrane-based polymer–ceramic nanocomposites generally does not possess the desired functional properties (electrical, dielectric, electronic, magnetic, optical, etc.). However, the introduction of a ceramic phase enhances the functional properties of nanocomposites.

4. Possible Applications

Even though membrane-based polymer–ceramic nanocomposites are still at an early stage of development, they may have great potential in many different applications because they possess properties of both the polymer membrane and the nanometer-sized ceramics. Some interesting areas are catalysis, gas separation, and filtration. Consider gas separation, for example. Traditionally gas separation and filtration take advantage of the permeability of membranes and the pore size of particles. By introducing ceramic nanoparticles into a polymer membrane, special functionality may be added to the resulting nanocomposite. For example, if the ceramic phase is a magnetic material such as γ-Fe_2O_3 or an oxide of another iron group metal, such as cobalt and nickel, a magnetic nanocomposite could be prepared. Such a nanocomposite may strongly bind gas molecules and can be used in gas separation or filtration for, as example, environmental protection.

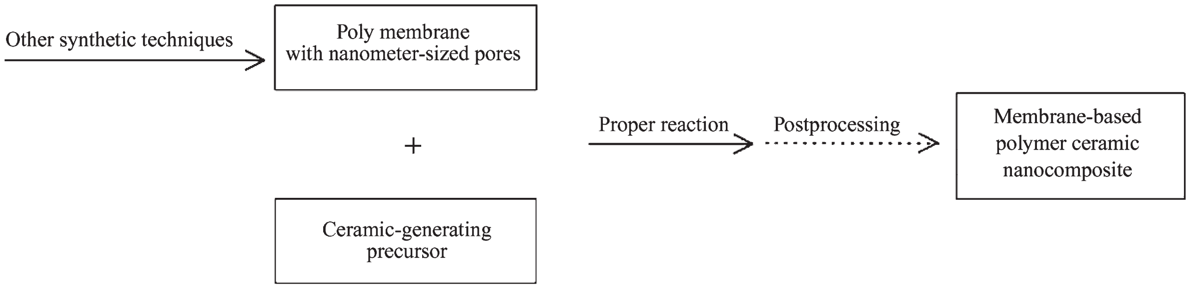

Other synthetic techniques
Poly membrane with nanometer-sized pores
+
Ceramic-generating precursor
Proper reaction
Postprocessing
Membrane-based polymer ceramic nanocomposite

Figure 3
Proposed alternative synthetic strategy for the preparation of membrane-based polymer–ceramic nanocomposites.

Functional materials constitute an unlimited and ever-expanding list. They include nanocomposites with good flame-retardant performance, even though the polymer is flammable (Dagani 1999), nanocomposites with high biodegradability, nanocomposites for lithography (nanolithography) (Whitesides 1997), and nanocomposites for electronics (nanoelectronics) (Murray *et al.* 1993).

5. *Future Directions*

Significant advances have been made in research on membrane-based polymer–ceramic nanocomposites, but there are technical difficulties remaining to be overcome and many new fields to be explored. Some brief discussions are presented for the reader's interest.

Of the remaining technical difficulties, one is the uneven distribution of nanoparticles across the polymer membrane cross-section—the distribution profile discussed above. The profile may have negative effects on the physical properties of composites. It is likely that an uneven distribution causes internal strain, which, in turn, results in inferior mechanical properties. It will be interesting and beneficial to disperse the ceramic particles evenly in the bulk of nanocomposites. This could be achieved by the use of very thin membranes, increasing the compatibility between the ceramic precursor and the polymer membrane, or taking a non-sol–gel synthetic approach.

Another technical difficulty is the limited number of ceramic materials available for the preparation of membrane-based polymer–ceramic nanocomposites through the *in situ* sol–gel reaction. Oxides of aluminum, silicon, titanium, and zirconium have been the most widely researched ceramic materials because the alkoxides of these elements are readily available. These ceramic materials may provide enormous opportunities in nanocomposites, but they may not be able to play a satisfactory role when demands on magnetic, electronic, dielectric, optic, or other functionalities arise. New synthetic techniques need to be developed to make membrane-based polymer–ceramic nanocomposites more attractive.

Regarding new fields to be explored, one is the development of new methods of preparing membrane-based polymer–ceramic nanocomposites. Figure 3 presents such a strategy. Unlike conventional methods, this strategy, although not yet researched, first creates polymer membranes with nanometer-sized pores, then introduces a ceramic precursor into the pores. It uses these pores as reaction vessels and forces ceramic materials to be generated inside the pores. In this way, it can be guaranteed that the ceramic particles will be nanometer sized. The best part of this strategy is that it can take advantage of many reaction methods, not solely the *in situ* sol–gel reaction. This strategy appears more promising with the advances in research on making polymer membranes with nanometer-sized pores (Martin 1996, Nguyen *et al.* 1999).

Another potential field may be membrane-based polymer–ceramic nanocomposites with a biopolymer as the organic phase. Biopolymers typically have low mechanical strength, and their reinforcement by the traditional powder technique is not suitable in many situations. It is anticipated that biopolymer–ceramic nanocomposites will be very promising in a variety of medical applications.

Of course, there are many other challenges remaining for membrane-based polymer–ceramic nanocomposites. In addition to developing better and more efficient synthetic methods and exploring new fields, searching for powerful characterization techniques and the ability to tailor the nanocomposites to application requirements will certainly be required.

Bibliography

Apichatachutapan W, Moore R B, Mauritz K A 1996 Asymmetric Nafion/(zirconium oxide) hybrid membranes via the *in-situ* sol–gel chemistry. *J. Appl. Polym Sci.* **62**, 417–26

Brinker C J, Scherer G W 1990 *Sol–Gel Science: The Physics and Chemistry of Sol–Gel Processing*. Academic Press, Boston

Dagani R 1999 Putting the "nano" into composites. *Chem. Eng. News* **77**, 23, 25–37

Deng Q, Moore R B, Mauritz K A 1998a Nafion$^{\circledR}$/(SiO2, ORMOSIL, and dimethylsiloxane) hybrids via *in situ* sol–gel reactions: characterization of fundamental properties. *J. Appl. Polym. Sci.* **68**, 747–63

Deng Q, Wilkie C A, Moore R B, Mauritz K A 1998b TGA-FTIR Investigation of the thermal degradation of Nafion® and Nafion®/[silicon oxide]-based nanocomposites. *Polymer* **39**, 5961–72

Gummaraju R V, Moore R B, Mauritz K A 1996 Asymmetric [Nafion®]/[Silicon oxide] hybrid membranes via the *in-situ* sol–gel reaction for tetraethoxysilane. *J. Polym. Sci. B Polym. Phys.* **34**, 2382–92

Hadjipanayis G C, Siegel R W (eds.) 1993 *Nanophase Materials: Synthesis–Properties–Applications*. Kluwer, Boston

Hench L L, West J K 1990 The sol–gel process. *Chem. Rev.* **90**, 1, 33–1, 72

Kunz M S, Shull K R, Kellock A J 1993 Colloidal gold dispersions in polymeric matrices. *J. Coll. Interface Sci.* **156**, 240

Lerner M M, Oriakhi C O 1997 In: Goldstein A N (ed.) *Handbook of Nanophase Materials* Dekker, New York

Mark J E, Lee C Y-C, Bianconi P A 1995 *Hybrid Organic–Inorganic Composites*. American Chemical Society, Washington, DC

Martin C R 1996 Membrane-based synthesis of nanomaterials. *Chem. Mater.* **8**, 1739–46

Mountz D A, Reuschle D A, Brister L B, Storey R F, Mauritz K A 1998 Swelling of and sol–gel reactions within poly (styrene-*b*-isobutylene-*b*-styrene) ionomers. *Polym. Prepr.* **39**, 1, 383– 4

Murray C B, Norris D J, Bawendi M G 1993 Synthesis and characterization of nearly monodisperse CdE (E = S, Se, Te) semiconductor nanocrystallites. *J. Am. Chem. Soc.* **115**, 8706–15

Nguyen C V, Beyers R B, Hawker C J, Hedrick J L, Jaffe R L, Miller R D, Remenar J F, Rhee H-W, Toney M F, Trollsas M, Yoon D Y 1999 Structure–property relationships for nanoporous poly(methyl-silsesquioxane) films with low dielectric constants prepared via organic–inorganic polymer hybrids. *Polym. Prepr.* **40**, 1, 398–9

Roecher A, Moller M 1995 Extraction of aqueous sol–gels with polymers. *Adv. Mater.* **7**, 151

Shao P L, Mauritz K A, Moore R B 1995 [Perfluorosulfonate ionomer]/[mixed inorganic oxide] nanocomposites via polymer–*in situ* sol–gel chemistry. *Chem. Mater.* **7**, 192–200

Siuzdak D A, Mauritz K A 1999 Surlyn®/[silicon oxide] hybrid materials. 2. Physical properties characterization. *J Polym. Sci. B Polym. Phys.* **37**, 143–54

Spatz J P, Roecher A, Moller M 1996 Golo nanoparticles in micellar polymer. *Adv. Mater.* **8**, 3220

Start P, Mauritz K A 1998 Morphology of sol–gel-derived Surlyn®/[silicon oxide] hybrids. *Polym. Prepr.* **39**, 1, 375–6

Wen J Y, Mark J E 1995a Sol–gel preparation of composites of poly(dimethylsiloxane) with SiO_2 and SiO_2/TiO_2 and their mechanical properties. *Polym. J.* **27**, 492–502

Wen J, Mark J E 1995b Synthesis, structure and properties of poly(dimethylsiloxanes) networks reinforced by *in-situ* precipitation of silica–titania, silica–zirconia and silica–alumina mixed oxides. *J. Appl. Polym. Sci.* **58**, 1135–45

Whitesides G M 1997 In: Siegel R W, Hu E, Roco M C (eds.) *R & D Status and Trends in Nanostructured Material, and Nanodevices in the United States*. International Technology Research Institute, Baltimore, MD, pp. 76–8

Young S K, Mauritz K A 1998 Structure–property relationships for novel Nafion®/ORMOSIL nanocomposites *Polym. Prepr.* **39**, 1, 377–8

Z. Pu
Cincinnati Financial Corporation, Ohio, USA

Mesoporous Molecular Sieves

In 1992 Mobil scientists discovered that ionic surfactants could be used as templates for the synthesis of a family of mesoporous siliceous materials which are now commonly referred to as M41S (Kresge *et al.* 1992). These have pore sizes in the range 25–100 Å, and a narrow pore size distribution. The ordered structure precipitates out of the synthesis mixture into a composite silica/organic structured phase, which is then calcined to remove the organic material. The resulting M41S materials contain regular arrays of uniform channels, the dimensions of which can be tailored through the choice of surfactants, auxiliary chemicals, and reaction conditions. The syntheses of mesoporous materials are generally thought to be controlled via a supramolecular templating mechanism and for their formation the so-called liquid crystal templating (LCT) mechanism is postulated, which proceeds via co-precipitation of the organic template with the silica species (see Fig. 1).

A number of excellent reviews have already been written about LCT of mesoporous materials and related areas of research (Corma 1997).

Mesoporous materials have potential catalytic/separation technology applications in the fine chemicals and pharmaceutical industries. Because of their larger pore size, these materials are not restricted to the catalysis of reactions involving small molecules, which is the case for zeolites.

The most studied of the mesoporous materials is MCM-41, which possesses a hexagonal array of uniform mesopores. The calcined mesoporous products have surface areas often greater than $1000\,\mathrm{m}^2\,\mathrm{g}^{-1}$, and exceptionally high sorption capacities.

Comprehensive accounts describing the synthesis and characterization of MCM-41 using various chain length alkyl trimethyl ammonium surfactants as templates are readily available in the literature. Characterization has been carried out by a variety of X-ray and neutron scattering experiments, high resolution and cryogenic transmission electron microscopy (TEM), electron diffraction, scanning electron microscopy (SEM), atomic force microscopy (AFM), sorption measurements, and solid-state NMR.

Studies on monitoring the effects of altering surfactant-silica ratios in M41S synthesis gels have led to the isolation of other structures with cubic and lamellar geometries. The lamellar product, MCM-50, was found to be unstable to calcination. The cubic M41S material was named MCM-48 and had Ia3d symmetry (by X-ray diffraction and TEM).

Ko *et al.* (1996) describe an interesting means of confirming the connectivity of the pores within mesoporous materials, by incorporating platinum wires into the mesopores as contrast agents.

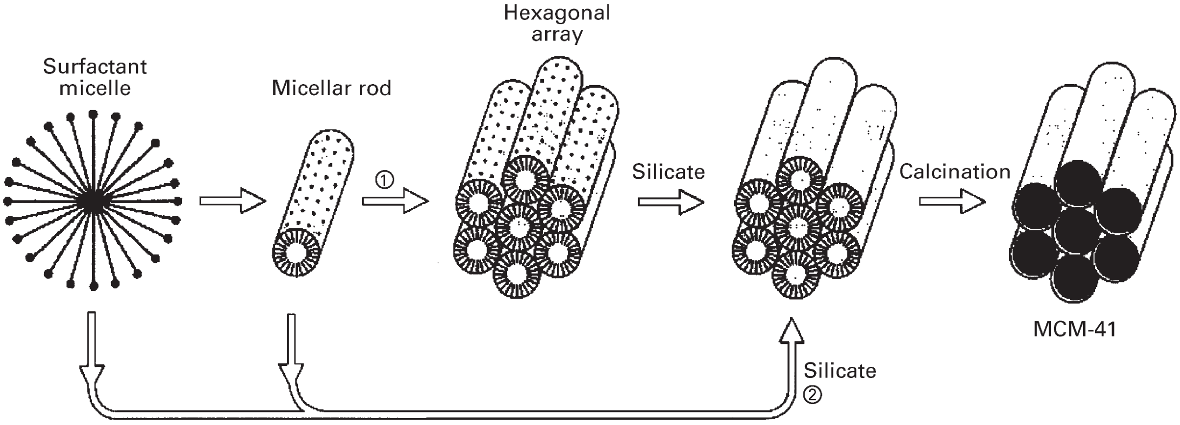

Hexagonal array
Surfactant micelle
Micellar rod
①
Silicate
Calcination
MCM-41
Silicate
②

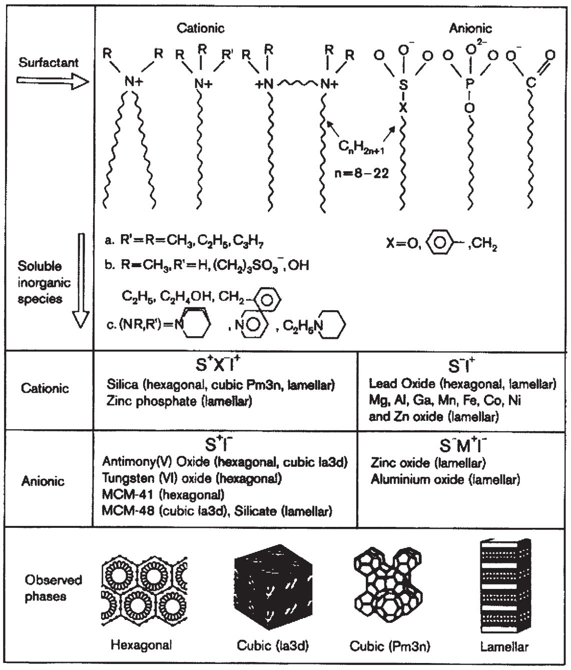
Figure 1
Possible mechanistic pathways for the formation of MCM-41: (1) liquid crystal phase initiated and (2) silicate anion initiated.

One main characteristic of all M41S materials is that they are thermodynamically metastable. The metastability of the MCM-41 silica walls is highlighted by work conducted by Kloetstra *et al.* (1997) who, on recrystallizing MCM-41 and MCM-48, produced mesoporous materials containing small amounts of framework ZSM-5, or other higher stability silicate species.

M41S materials have been synthesized in both basic and acidic media, and all the variables (such as pH, gel composition, and temperature) in the process have been thoroughly investigated by various research groups. Complete phase diagrams for the M41S synthesis gels are now available in the literature.

Control of the rate of silicate polymerization by using different pH conditions and controlling the aluminosilicate reactivity in the synthesis gels produces materials with thicker pore walls and improved thermal and hydrothermal stabilities. The effect of gel aging time, temperature, stirring, source of silica, and alkyltrimethylammonium template (CTAB or CTACl), and overall concentration of the reactants on the quality of the MCM-41 product was studied by Cheng *et al.* (1997). It was found that highly basic gels, or gels containing low silicate concentrations, favored lamellar products.

Further work was carried out by Sayari *et al.* (1997) where hydrothermal restructuring of MCM-41 was used to produce large pore MCM-41 products. It was shown that the pore size enlargement was accompanied by a significant improvement of pore size uniformity, a gradual decrease of the specific surface area, and the pore wall thickening. It was suggested that the post synthetic hydrothermal restructuring is a convenient synthesis route for MCM-41 silicates with a bimodal pore size distribution with controllable amounts of microporosity.


Surfactant
Cationic
Anionic
R
R'
N+
+N
S
X
P
O
C
C_nH_{2n+1}
n=8–22
Soluble inorganic species
a. $R'=R=CH_3, C_2H_5, C_3H_7$
b. $R=CH_3, R'=H, (CH_2)_3SO_3^-, OH$
C_2H_5, C_2H_4OH, CH_2-
c. $(NR,R')=N$
, N
, C_2H_5N
X=O, ,CH_2
Observed phases
Hexagonal
Cubic (Ia3d)
Cubic (Pm3n)
Lamellar

Cationic	$S^+X^-I^+$ Silica (hexagonal, cubic Pm3n, lamellar) Zinc phosphate (lamellar)	S^-I^+ Lead Oxide (hexagonal, lamellar) Mg, Al, Ga, Mn, Fe, Co, Ni and Zn oxide (lamellar)
Anionic	S^+I^- Antimony(V) Oxide (hexagonal, cubic Ia3d) Tungsten (VI) oxide (hexagonal) MCM-41 (hexagonal) MCM-48 (cubic Ia3d), Silicate (lamellar)	$S^-M^+I^-$ Zinc oxide (lamellar) Aluminium oxide (lamellar)

Figure 2
Synthesis pathways using different combinations of surfactant and soluble inorganic species. S is the surfactant, X is a halide ion, M is a metal ion, and I is the inorganic ionic precursor (Huo *et al.* 1994).

Small particles of MCM-41 are obtained if reaction times are kept short. Cutting the reaction times short can, however, jeopardize the silica condensation process leading to poorly polymerized products.

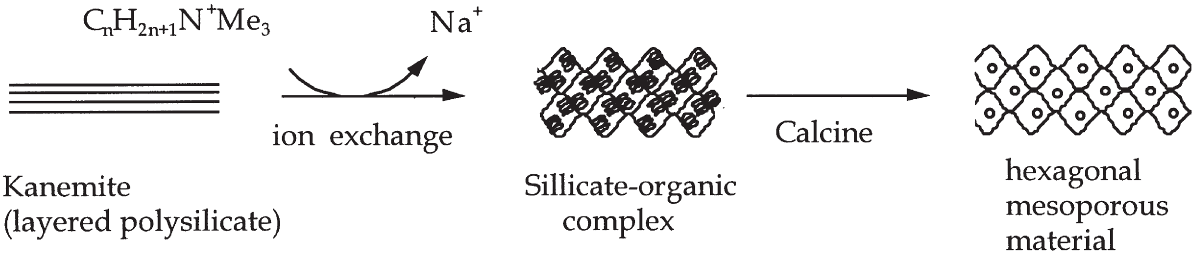


Figure 3
Scheme showing the mechanism of formation of hexagonal mesoporous FSM-16 from kanemite.

Microwave heating overcomes this problem by speeding up the condensation step.

Large headgroup cationic surfactants such as $C_{16}H_{33}N(C_2H_5)_3Br$ with acid or base synthesis conditions lead to three-dimensionally structured silica products containing cage-like voids, with dimensions somewhat larger than the channel diameters measured for the M41S materials. SBA-1 is a cubic cage mesostructure with $Pm\bar{3}n$ space group symmetry, and SBA-2 is a three-dimensional hexagonal cage structure with $P6_3/mmc$ symmetry (Huo *et al.* 1994).

1. Neutral Templating Routes

A dilute regime method for the synthesis of mesoporous solids was reported where the template was a nonionic surfactant (a primary amine or polyethylene oxide) and the silica precursors were alkoxides such as tetraethoxysilane (TEOS) (Tanev and Pinnavaia 1995, Bagshaw *et al.* 1995). The mesoporous products resemble M41S materials, but have slightly less long-range order and have thicker silicate walls, smaller crystallite size, and improved textural mesoporosity (improved access to framework pores). The neutral synthesis system has the added advantage that the surfactant could be removed by neutral solvent extraction, avoiding the calcination step.

A neutral alkoxide/polyoxyethylene templating route can also be used to produce mesoporous products with the pores arranged in random worm-like motifs, with no distinguishable hexagonal or cubic long range ordering. The products, named MSU-X, are formed by firstly obtaining a homogeneous alkoxide/surfactant solution by aging at room temperature, and secondly initiating the mesophase assembly by adding fluoride ions. Fluoride ions had previously been shown to catalyze alkoxide hydrolysis and mesostructure cross-linking.

Similar types of material called TUD-1 can be made using triethanolamine as a template. This route offers great versatility and is easily up-scalable (Maschmeyer *et al.* 2000).

2. Gemini Surfactants as Templates

Dual chain dialkyldimethylammonium salts (gemini surfactants) can also be used as the templating agents. For gemini surfactants with the generalized molecular formula $C_{16}H_{33}(CH_3)_2[(CH_2)_nCH_3]N^+$, where $n = 1–12$, a range of structures can be produced. For odd n (1, 3, 5, 7), hexagonal MCM-41 was produced. For even *n* (2, 4, 6, 8) and any *n* greater than 7, the product was a lamellar phase.

Huo *et al.* used somewhat different gemini surfactants as templates for a whole range of mesoporous structures: MCM-41 (2D hexagonal, p6mm); MCM-48 (cubic, Ia3d); MCM-50 (lamellar); SBA-1 (cubic, Pm3n), SBA-2 (3D hexagonal, $P6_3/mmc$) and SBA-3 (2D hexagonal, p6mm). A new mesoporous supercage structure, SBA-2, which was analogous to a zeolite cage structure was also synthesized.

Virtually at the same time as M41S mesoporous silicas were first being synthesized, Inagaki *et al.* (1993) reported the synthesis of hexagonally packed channels from layered polysilicate kanemite. The mesoporous product was named FSM-16.

Kanemite consists of flexible, poorly polymerized silicate layers which buckle around the intercalated surfactant molecules. The kanemite-derived materials are more highly condensed, and exhibit thicker silicate walls than the M41S materials. Chen *et al.* (1995) suggest that the thermal and hydrothermal stability of FSM-16 is somewhat better than for MCM-41. Full characterizations of the kanemite-derived mesoporous silica product are described by Inagaki *et al.* (1996).

3. True Liquid Crystal Templating (TLCT) of Mesoporous Materials

Attard *et al.* successfully synthesized mesoporous silica from synthesis gels containing nonionic polyoxyethylene (POE) surfactants at concentrations required to form liquid crystal phases at room temperature (Attard *et al.* 1995). Hexagonal, cubic, and lamellar products were obtained, and the former two were stable to calcination.

BET nitrogen adsorption measurements showed that the hexagonal $C_{12}EO_8$-templated calcined product has a specific surface area greater than $1000\,m^2\,g^{-1}$, and a very narrow pore size distribution at 20 Å diameter. These values compared well with the large surface areas obtained for MCM-41, the LCT analogue.

Striking similarities were observed by polarizing microscopy between the birefringent patterns obtained for the hexagonal POE/water liquid crystal, the synthesis gel (after removal of the methanol hydrolysis product), the as-synthesized product and the calcined product. These observations strongly supported the TLCT theory, in which it is postulated that the surfactant forms a liquid crystal phase before the condensation of the silica, and remains in the same phase throughout the synthesis.

One main advantage of TLCT over the dilute regime M41S synthesis is that by using a preformed liquid crystal phase as the templating medium, the structure of the product is known *a priori* from the binary phase diagram of the surfactant/water system.

4. Macroscopic Morphology Control

The alkaline M41S synthesis procedure generally produces very fine particles of mesoporous material, which are suitable for catalysis applications but not for some specialized applications. Acidic conditions have become widely used for mesoporous silica synthesis because of the greater control which can be imposed on the macroscopic morphology of the product. The lack of ionic bonds at the silica/organic interfaces for acidic preparations allows more fluidity in the self-organizing system with respect to basic preparations. This effectively makes macroscopic morphology control possible in acidic systems.

For an acidic mesoporous M41S silica synthesis under aqueous conditions Yang *et al.* (1997) reported the formation of a remarkable array of macroscopic shapes, surface patterns and channel arrangements. The reaction conditions favored curved morphology (toroidal, disk-like, spiral, and spheroidal particles). Powder X-ray diffraction showed that the underlying mesostructure for all the morphologies was ordinary hexagonal MCM-41. The proposed initial step for the morphogenesis of these shapes was the formation of a silicate liquid-crystal embryo with hexagonal cross-section. This grew over time and depending on the various reaction conditions, was subject to different degrees of curvature (Yang *et al.* 1997). Many other examples exist where monolithic silica, films, and spheres are produced (Yang *et al.* 1996, Zhao *et al.* 1998).

5. Functionalization for Catalytic Applications

Much work has been directed towards the introduction of transition metals into mesoporous materials. Transition metal oxides have variable oxidation states, and can therefore be used in tailoring catalytic, electronic and magnetic properties of materials, comprehensive reviews are available covering this subject (Corma 1997).

Siliceous solids can be functionalized for catalysis applications by framework substitutions, or activated by post-synthetic surface derivatizations. The four possible methods of derivatization are:

(i) Primary modifications: direct grafting of active sites onto the surface silanol groups of calcined or solvent extracted materials;

(ii) Second and higher order modifications: grafting onto previously functionalized surfaces;

(iii) One-pot synthesis of functionalized mesoporous silica; or

(iv) Transformation of materials prepared under conditions (i)–(iii) by further treatment such as heating, evacuation, calcination and/or reduction.

Transition metals incorporated into the MCM-41 framework, or grafted to the MCM-41 surface include titanium (Maschmeyer *et al.* 1995), platinum, vanadium, manganese, boron, aluminium, zinc, zirconium, binary lanthanum oxides (Na/La, K/La, Rb/La, Cs/La), cobalt (Maschmeyer *et al.* 1997), ruthenium, nickel–molybdenum clusters, and tin–molybdenum clusters.

A novel means of introducing metal ions into mesoporous structures is to use surfactant templates with organometallic complex head groups. Kevan and co-workers used a surfactant with Cu^{II} ions coordinated to a macrocyclic cage headgroup to template the formation of MCM-41. Also surfactant templates with ruthenium complex headgroups for the TLCT of mesoporous silica have been used.

Other approaches to functionalizing mesoporous silica include the incorporation of heteropolyacids such as $H_3PW_{12}O_{40}$, into the mesopores of MCM-41 materials. Burkett *et al.* (1996) discovered that the co-condensation of TEOS with organosilanes in the presence of a surfactant template led to the synthesis of organically functionalized mesoporous silicas.

For TLCT systems one-pot synthesis of functionalized silica is in theory quite straightforward. The key criterion which determines whether a particular synthesis gel will successfully produce a mesoporous structured catalyst is simply that the addition of the functional moiety does not disturb the liquid crystal template.

6. Further Applications

Advanced separation/purification, electronic, optical and electron transfer applications have been suggested for MCM-41-type materials, generally relying on the special characteristics of species encapsulated in the material mesopores. Potential applications other than catalysis include:

(i) stationary phases for chromatography,
(ii) environmental remediation,
(iii) minireactors for electron transfer reactions,
(iv) mesoporous hosts for quantum confinement,
(v) mesoporous supports for molecular wires (Wu and Bein 1994),
(vi) shape-selective polymerization, and
(vii) protein synthesis.

Bibliography

Attard G S, Glyde J C, Goltner C G 1995 Liquid-crystalline phases as templates for the synthesis of mesoporous silica. *Nature* **378**, 366–8

Bagshaw S A, Prouzet E, Pinnavaia T J 1995 Templating of mesoporous molecular-sieves by nonionic polyethylene oxide surfactants. *Science* **269**, 1242–4

Burkett S L, Sims S D, Mann S 1996 Synthesis of hybrid inorganic–organic mesoporous silica by co-condensation of siloxane and organosiloxane precursors. *J. Chem. Soc. Chem. Commun*, 1367–8

Chen C Y, Xiao S Q, Davis M E 1995 Studies on ordered mesoporous materials. *Microporous Mater*. **4**, 1–20

Cheng C F, Park D H, Klinowski J C 1997 Optical parameters for the synthesis of the mesoporous molecular sieve [Si]-MCM-41. *J. Chem. Soc., Faraday Trans.* **93**, 193–7

Corma A 1997 From microporous to mesoporous molecular sieve materials and their use in catalysis. *Chem. Rev.* **97**, 2373–419

Huo Q S, Leon R, Petroff P M, Stucky G D 1995 Mesostructure design with gemini surfactants—supercage formation in a three-dimensional hexagonal array. *Science* **268**, 1324–7

Huo Q S, Margolese D I, Ciesla U, Feng P Y, Gier T E, Sieger P, Leon R, Petroff P M, Schuth F, Stucky G D 1994 Generalized synthesis of periodic surfactant in organic composite-materials. *Nature* **368**, 317–21

Inagaki S, Fukushima Y, Kuroda K 1993 Synthesis of highly ordered mesoporous materials from a layered polysilicate. *J. Chem. Soc., Chem. Commun.* 680–2

Inagaki S, Sakamoto Y, Fukushima Y, Terasaki O 1996 Pore wall of mesoporous molecular sieve devised from kanemite. *Chem. Mater*. **8**, 2089–95

Kloestra K R, van Bekkum H, Jansen J C 1997 Mesoporous material containing framework tetrasilicate by pore-wall recrystallisation. *J. Chem. Soc., Chem. Commun.* 2281–2

Ko C H, Ryoo R 1996 Imaging the channels in mesoporous molecular sieves with platinum. *J. Chem. Soc., Chem. Commun.* 2467–8

Kresge C T, Leonowicz M E, Roth W J, Vartuli J C, Beck J S 1992 Ordered mesoporous molecular-sieves synthesized by a liquid-crystal template mechanism. *Nature* **359**, 710–2

Maschmeyer T, Oldroyd R D, Sankar G, Thomas J M, Shannon I J, Klepetko J A, Masters A F, Beattie J K, Catlow C R A 1997 Designing a solid catalyst for the selective low-temperature oxidation of cyclohexane to cyclohexanone. *Angew. Chem. Int. Ed. Engl.* **36** (15), 1639–42

Maschmeyer T, Rey F, Sankar G, Thomas J M 1995 Heterogeneous catalysts obtained by grafting metallocene complexes onto mesoporous silica. *Nature* **378**, 159–62

Maschmeyer T, Shan Z, Jansen J C 2000 Inorganic oxides with mesoporosity or combined meso- and microporosity and process for preparation thereof. WO 01555 1A1 World Patent

Sayari A, Liu P, Kruk M, Jaronieo M 1997 Characterisation of large pore MCM-41 molecular sieves obtained via hydrothermal restructuring. *Chem. Mater*. **9**, 2499–506

Tanev P T, Pinnavaia Y J 1995 A neutral templating route to mesoporous molecular sieves. *Science* **267**, 865–7

Wu C G, Bein T 1994 Conducting polyaniline filaments in a mesoporous channel host. *Science* **264**, 1757–9

Yang H, Coombs N, Ozin G A 1997 Morphogenesis of shapes and surface patterns in mesoporous silica. *Nature* **386**, 692–5

Yang H, Kuperman A, Coombs N, Mamiche-Afara S, Ozin G A 1996 Synthesis of oriented films of mesoporous silica on mica. *Nature* **379**, 703–5

Zhao D Y, Yang P D, Hur Q S, Chmelka B F, Stucky G D 1998 Topological construction of mesoporous materials. *Curr. Opin. Solid-State and Mater. Sci.* **3**, 111–21

Th. Maschmeyer[a,b], M. E. Raimondi[b] and
J. M. Seddon[b]
[a] *Delft University of Technology, The Netherlands*
[b] *Imperial College of Science, Technology and Medicine, London, UK*

Metal Matrix Composites with Roughly Equiaxed Reinforcements: Microstructure and Mechanical Behavior

The early developmental work on metal matrix composites concentrated on continuous fiber reinforcement because this potentially gives the maximum improvement in elastic modulus and strength (see *Composites, Physical Properties of**). However, this enhancement in mechanical properties was obtained at considerable cost, resulting from the high price of strong, continuous fibers, and the difficulty in fabricating such composites on a commercial scale. These limitations led to the development of discontinuously reinforced metal matrix composites containing roughly equiaxed reinforcing particles, and while these composites do not attain the very highest mechanical property enhancement of continuous fiber-reinforced metals, they still provide sufficient improvement to make them commercially attractive in some applications. Discontinuously reinforced composites can be made by a variety of methods, many of which are modifications of normal metal processing methods, resulting in much lower processing costs than for continuously reinforced composites (see *Metal Matrix Composites: Matrices and Processing**). In addition, the discontinuous reinforcement is readily available in the form of low-cost ceramic powders, such as the silicon carbide (SiC) and alumina (Al_2O_3) grits used in the grinding and polishing industry. Since the major applications take advantage of composites with low weight, high stiffness, and high strength, the major developments have involved low-density metals, particularly aluminum and magnesium, as the matrix of the composite.

1. *The Reinforcement*

Since the object of the reinforcement is to provide increased stiffness and strength to the unreinforced matrix, ceramic particles with their large elastic modulus and high strength are ideal as the reinforcing particles. In principle, many ceramic powders meet these requirements, but there are other restrictions that influence the choice of the reinforcement for a particular composite. Many of the ceramic particles of interest are thermodynamically unstable when they are in contact with pure metals, and will react to form reaction compounds at the interface between the particles and the surrounding matrix. Reaction compound layers at the matrix–reinforcement interface can degrade the interfacial strength of the composite, resulting in deterioration in the mechanical properties (Metcalfe 1974, Lloyd *et al.* 1990). The kinetics of the interfacial reaction will increase with increasing temperature, so the temperature history of the composite, either during processing or during service, has to be considered in choosing the appropriate matrix–reinforcement combination. When aluminum is the matrix, there are very few ceramics that are stable. SiC will react with molten aluminum to form aluminum carbide, Al_4C_3:

$$4Al + 3SiC \leftrightarrow Al_4C_3 + 3Si \quad (1)$$

Aluminum carbide is very brittle, and very reactive in moist environments, so its formation degrades both the mechanical and corrosion properties of the composite. The reaction is accompanied by an increase in silicon level of the surrounding matrix, which can result in a change in the properties of the matrix alloy. Fortunately, SiC is stable in most aluminum alloys in the solid state, so reactivity is not a consideration when these composites are made without exposure to liquid aluminum. This is not the case with other reinforcements, and aluminum oxide, Al_2O_3, particles are an interesting example in this respect. Al_2O_3 is thermodynamically stable in pure aluminum, but commercial alloys often contain elemental additions that can react with Al_2O_3, in particular, magnesium. The reaction equations in this case are

$$3Mg + Al_2O_3 \leftrightarrow 3MgO + 2Al \quad (2)$$

and

$$3Mg + 4Al_2O_3 \leftrightarrow 3MgAl_2O_4 + 2Al \quad (3)$$

The first reaction (Eqn. (2)) to form MgO predominates when there is more than about 1 wt.% magnesium in the alloy, while spinel, $MgAl_2O_4$, will form in alloys containing lower magnesium contents, and can form in the solid state. It is obvious from Eqns. (2) and (3) that aluminum oxide is not a good reinforcement for magnesium alloys, but ceramic carbide particles are a good choice because magnesium does not have a stable carbide. Reaction stability is one of the most important factors in choosing the appropriate reinforcement, and in some cases attempts have been made to provide protective coatings on the surface of the reinforcement to avoid interface reaction (see *Composites: Interfaces**).

There are other factors that come into consideration, in addition to reaction stability, when choosing an appropriate reinforcement. Ceramic powders come in a wide range of sizes, but most composites use particles in the 5–25 μm size range. The larger the ceramic particle the lower is its fracture strength. This is because ceramic particles are brittle solids, where fracture is initiated by critical sized flaws present in the particles, and the probability of a particle containing a critical flaw increases with increasing particle size. Premature failure of the reinforcing particles will degrade the composite properties. The method of composite manufacture (see *Metal Matrix Composites: Matrices and Processing**) can also influences the size of reinforcing particles. If the composite is made by blending a mixture of solid powers, as in powder metallurgy manufacturing methods, the ratio of the metal powder size to that of the reinforcing particle is important if a uniform distribution of the reinforcement is to be achieved. In general, the more similar the size ranges of the metal and ceramic particles, the more uniform the reinforcement distribution. The situation is somewhat different if a molten metal mixing route is used for incorporating the particles into the matrix. In this case the ceramic particles are stirred into a liquid melt, and the viscosity of the resulting composite melt is important. If the ceramic particles are too fine, the viscosity of the melt increases to a level where it can no longer be poured from the mixing container. For molten metal mixing processes the reinforcing particles need to be in the 10–20 μm size range. Finally, the cost of the ceramic particles is important, since they are a major component of the cost of the final composite.

Considering all these factors, SiC and Al_2O_3 have emerged as the major reinforcements for particle-reinforced composites.

2. *Composite Production and Fabrication*

There are a variety of methods for producing composites (see *Metal Matrix Composites: Matrices and Processing**; Lloyd 1997), but they are of two generic types: powder metallurgy and molten metal processing. The powder process is essentially a solid-state processing route, involving relatively low temperatures, so reactivity problems between the matrix and reinforcement are significantly reduced. Since the process involves blending of the alloy matrix and ceramic powders, there is little limitation on the volume fraction of reinforcement that can be produced, which is important if very high mechanical properties, such as elastic modulus, are required. However, to obtain optimum properties a good interfacial bond has to be developed between the particles, both metal–metal and metal–ceramic. This requires extensive secondary deformation after the initial consolidation of the blended powders, and this is usually obtained by extrusion of the initial, compacted billet. The initial billet has relatively low mechanical properties because the compaction of the powder to a billet with about 90–95% density does not introduce sufficient deformation to break up the oxide skin on the surface of the metal powders, which is required to obtain a good metal-to-metal bond. The initial billet has to be reduced in section by a ratio of about 20:1 to achieve optimum interparticle bonding and good mechanical properties. The molten metal processing routes do not have this limitation, and they can produce directly a composite with good mechanical properties. For composites made by molten metal mixing, the liquid composite can be cast directly to final shape (Cox *et al.* 1993), provided the viscosity of the melt is not too high to inhibit mold filling. The viscosity of the melt increases with increasing volume fraction of reinforcing particles, so this processing route is limited to reinforcement levels of about 20–25 vol.%. Other molten metal processing routes, such as melt infiltration into a ceramic powder bed, and some spray deposition methods are not limited by viscosity considerations. However, since they all involve molten metal in contact with a ceramic powder, the reactivity issues discussed above have to be taken into consideration.

One of the major attributes of particle-reinforced metal matrix composites is that they can be fabricated by conventional shaping processes, such as extrusion and rolling, into a variety of final forms, including sheet, bar, and hollow sections (Dixon and Lloyd 1996). In addition to developing the final form, the extrusion or rolling has an important effect on the distribution of the reinforcing particles. Prior to the shaping process the particles tend to be quite inhomogeneously distributed. In the molten processing methods, the particles are located in a network surrounding the dendrites formed during solidification. As the melt solidifies the dendrites grow and push the particles ahead of them, so that when solidification is complete the particles are inhomogeneously distributed in the interdendritic regions (Fig. 1(a)). The deformation accompanying the subsequent extrusion or rolling redistributes the particles more homogeneously within the metal matrix, although on a local scale the particles are still clustered to some extent (Fig. 1(b)). A similar situation occurs with the solid-state processing routes, but in this case the initial reinforcing particle distribution reflects the efficiency of the blending process and the relative size difference between the metal and ceramic powders. Some care has to be taken during the shaping process because the deformation involved is quite extensive, and high stresses can develop within the brittle ceramic particles, causing them to crack and become damaged. To minimize this effect the forming is usually carried out at elevated temperatures to minimize the stresses involved.

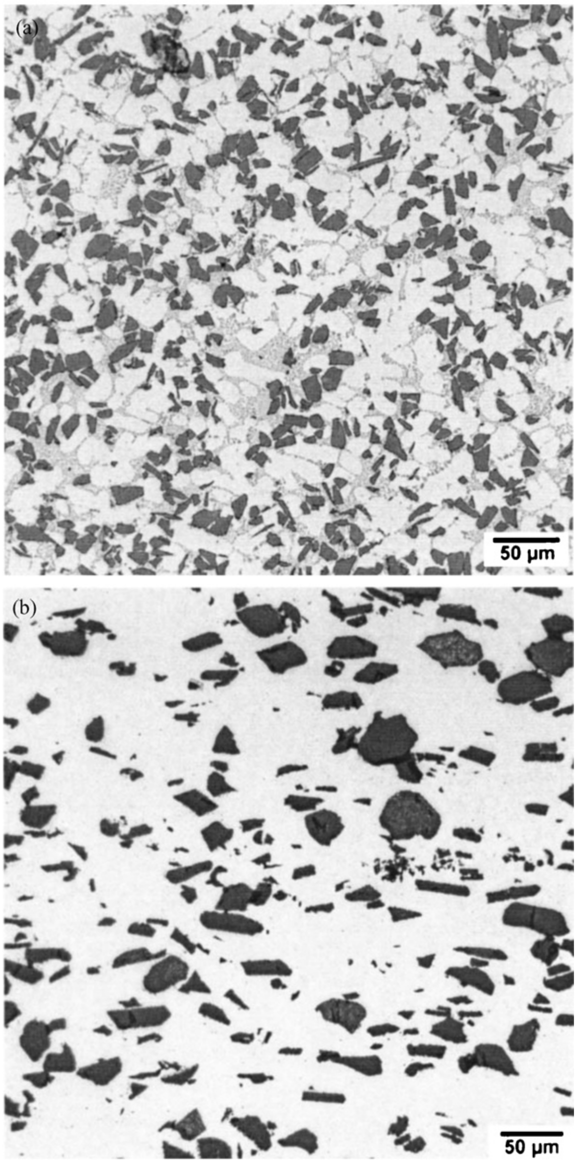


Figure 1
Dendritic microstructure (a) at solidification and (b) after processing.

When aluminum is used commercially in the unreinforced form, it contains other elemental additions in a range of alloys that provide different combinations of physical and mechanical properties. In addition, some of these alloys can be given different heat treatments to further extend the strength beyond that obtained by composition alone. Of particular importance are the Al–Mg–Si, Al–Cu, and Al–Mg–Zn alloys that can be heat treated to provide additional strength through precipitation hardening, often referred to as age hardening. Any of the commercially available alloys can be used for the composite matrix, provided the appropriate combination of reinforcement and processing route is used to minimize interfacial reaction.

3. Mechanical Properties

When the complete range of alloy compositions and heat treatments (or tempers) in aluminum alloys is considered, the strength range extends from about 20 MPa to over 600 MPa (Hatch 1984). This is one of the main reasons why aluminum alloys are used in such a wide range of products, from household foil to aircraft structures. However, while the strength range is very large, the elastic modulus is much more limited, extending from around 69 GPa to 72 GPa in the commercially available alloys. It is with regard to improving the elastic modulus that composites make their major impact on the properties of metals in general, and aluminum and magnesium alloys in particular. Other mechanical properties are also affected by the addition of reinforcing particles, and not always in a positive fashion. Unlike continuous fiber-reinforced composites, particle-reinforced materials have quite isotropic mechanical properties.

3.1 Elastic Properties

The elastic modulus is a global property that is largely independent of the distribution of reinforcing particles, but it is dependent on the quality of the particle–matrix interface, which has to have adequate strength to transfer loading in the matrix to loading the particle. The measured values in the literature show a wide degree of scatter, which reflects different methods of measurement—dynamic measuring methods tend to give larger values than static measurements—and the quality of the composite. The quality of the composite will involve factors such as the interfacial strength, porosity, which is often associated with "unwet" particle clusters, and damaged particles. As noted above, particles can be cracked during fabrication, reducing their ability to carry stress, which will reduce the elastic modulus (measurement of the elastic modulus has been used to monitor particle fracture with progressive deformation (Lloyd (1991)). The elastic modulus can also be affected by the presence of residual stresses in the composite, which are often generated during fabrication as a result of the coefficient of expansion mismatch between the alloy matrix and the reinforcing particles.

The simplest approach to representing the elastic modulus is the upper (K_U) and lower (K_L) bounds for the bulk modulus:

$$K_L = K_M + \frac{V_P}{1/(K_P - K_M) + 3(1 - V_P)(3K_M + 4G_M)} \tag{4a}$$

and

$$K_U = K_P + \frac{1 - V_P}{1/(K_M - K_P) + 3V_P/(3K_P + 4G_P)} \tag{4b}$$

where the subscripts M and P refer to matrix and particle properties, respectively, V_P is the volume fraction of reinforcing particles, and G is the shear modulus (see *Composites, Physical Properties of**). Young's modulus, E, can then be obtained using the standard relationship $K = E/3(1-2\nu)$, where ν is Poisson's ratio. These bounds are quite broad, taking no account of particle shape, and Halpin (1984) used a micromechanics approach to take account of the particle aspect ratio:

$$E_C = \frac{E_M(1 + 2sqV_P)}{1 - qV_P} \tag{5}$$

where

$$q = \frac{(E_P/E_M - 1)}{(E_P/E_M + 2s)} \tag{6}$$

E_C is the Young's modulus of the composite, and s is the particle aspect ratio, defined as the ratio of largest to smallest perpendicular particle cross-sectional radii.

An alternative approach, which also takes account of the particle shape, is to use the Eshelby equivalent inclusion method (Clyne and Withers 1993).

Figure 2 shows some experimental upper and lower bounds for the Young's modulus of aluminum reinforced with silicon carbide particles. The results fall within the range predicted by Eqn. (4), and give reasonable agreement with Eqn. (5) and the Eshelby approach for $s \cong 1.5$, which is reasonable for the reinforcements involved.

A point regarding the modulus which is often overlooked, but which is commercially important, is that the modulus improvement given by composites is maintained, to essentially the same extent, at elevated temperatures. So the decrease in elastic modulus with increasing temperature reflects the decrease in the modulus of the matrix (Lloyd *et al.* 1994).

3.2 Yield Strength

Despite many reports to the contrary in the literature, the addition of particle reinforcement produces very little real increase in the yield strength over that of the matrix alloy. The confusion in the literature arises because of the usual method of measuring the yield strength in unreinforced material, and the nature of

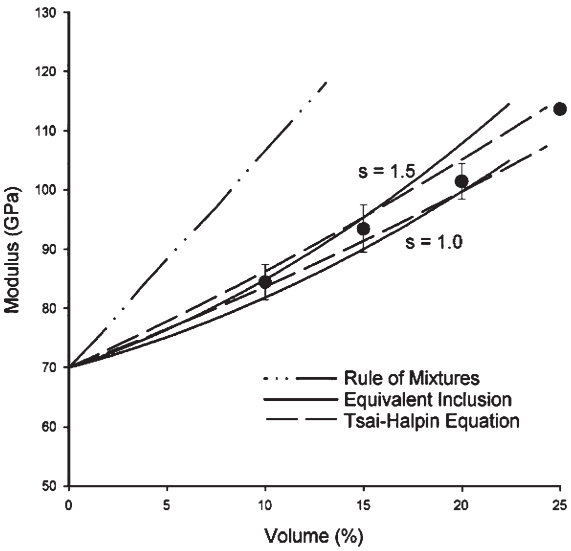


Figure 2
Young's modulus of SiC particle-reinforced aluminum composites. Filled circles are experimental values; lines correspond to different models.

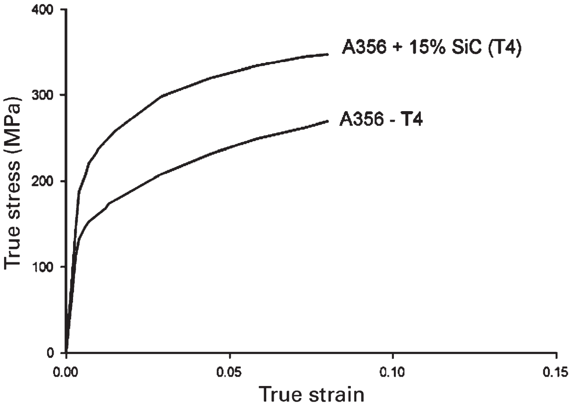


Figure 3
Stress–strain curves of a composite of Al–Si–Mg casting alloy A356 reinforced with 15% SiC particles in the as-quenched condition (T4) compared with that of the unreinforced matrix.

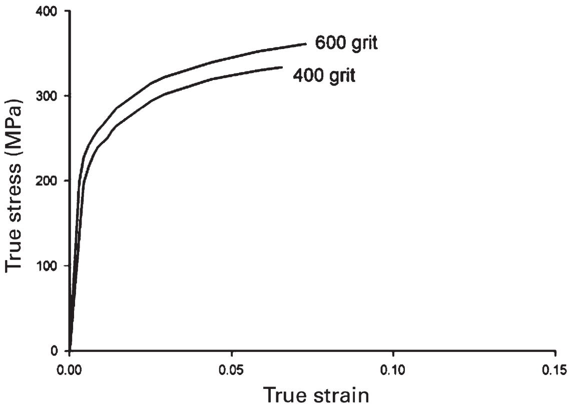


Figure 4
Stress–strain curves of composites of Al–Si–Mg casting alloy A356 reinforced with 20% SiC particles in the as-quenched condition (T4) as a function of particle size (600 grit size and 400 grit size correspond to roughly 9 μm and 17 μm average particle diameter, respectively).

the elastic-to-plastic transition in composites. The yield strength in unreinforced material is defined as the stress at which a plastic strain of 0.2% is achieved. In unreinforced metals the transition from elastic to plastic behavior is relatively abrupt, and the rate of work hardening decreases quite rapidly. In composites the elastic-to-plastic transition is much more gradual, and the work hardening rate is much higher at low strains. Figure 3 shows this difference in behavior, and it results in an apparent increase in the yield stress. At larger strains the increased initial work hardening rate gives a composite stress–strain curve above that of the matrix alloy, and a higher ultimate strength. The higher work hardening rate is a result of several factors. (i) There is locally a high dislocation density associated with reinforcing particles that results from the misfit in the coefficient of thermal expansion. This causes a higher flow stress overall, as well as a higher rate of dislocation accumulation with continued straining, resulting from an increased density in dislocation sources. (ii) The influence of the constraints the elastic reinforcing particles have on the plastic deformation of the matrix, owing to elastic misfit between the particle and the matrix; this can be modeled using the Eshelby approach or using finite element models. (iii) The way dislocations rearrange themselves with increasing strain, which lowers the total energy of the deformed state and reduces the resulting work hardening rate, can be impeded in the presence of elastic particles. All of these factors have been considered in the literature (Withers *et al.* 1989, Taya *et al.* 1991, Clyne and Withers 1993), but the models are dependent on a range of microstructural features, such as particle shape, size, and distribution, which makes analysis difficult. Experimentally, refinement of particle size, for reinforcement volume fractions up to about 25 vol.%, also results in an increased work hardening rate, and an apparent increase in strength, as shown in Fig. 4.

When a heat-treatable alloy is used for the matrix, the composite can be aged at elevated temperature to increase the strength of the matrix, and hence the strength of the composite. In some cases the composite may age harden somewhat more quickly than the unreinforced alloy, so the optimum heat treatment to obtain the appropriate temper properties may differ from that of the unreinforced alloy. Differences in aging kinetics are attributed to the presence of lattice defects generated in processing the composite (Prangnell and Stobbs 1991).

The yield strength of the composites decreases with increasing test temperature, which mainly reflects the temperature dependence of the matrix flow stress.

3.3 Tensile Elongation

The increase in elastic modulus and ultimate strength in particle-reinforced composites is accompanied by a decrease in the tensile elongation, or ductility, which is a major limitation to the use of these materials. Many factors can influence the ductility of the composite, but the major ones are the matrix strength–ductility dependence, the reinforcing particle distribution, and the particle–matrix interfacial strength (Lloyd 1995). The importance of matrix ductility is because of the fact that any global plastic strain to the composite results in much higher local matrix strain since the ceramic particles can only deform elastically. Particle distribution is very important because in regions where the particles are closely spaced, as in particle clusters, they constrain the plastic flow of the adjacent matrix, resulting in locally high complex stresses that result in premature failure. This failure may occur in the matrix regions between the particles if the matrix fracture stress is exceeded, or at the particle–matrix interface if the interfacial strength is exceeded. A high interfacial strength is obviously a requirement for optimum mechanical properties with such composites. Other factors, such as reinforcement size, where very large particles show particle cracking, and particle shape, where irregular shaped particles can give rise to local stress concentrations, can have an impact on the composite ductility. However, these tend to be second-order effects in good-quality composites.

3.4 Fracture Toughness

The fracture toughness to some extent mirrors the tensile ductility, decreasing with increasing volume fraction of reinforcement. However, the toughness decreases most rapidly with the first addition of reinforcing particles, and decreases only slightly for reinforcement loading beyond about 10 vol.% (Mortensen 1990, Davidson 1991). The toughness is much less sensitive to the degree of aging (i.e., the matrix strength) than to the tensile elongation, and there are some data suggesting that coarse particles are beneficial compared with smaller ones. Typical toughness values are in the 15–20 $\mathrm{MNm}^{-3/2}$ range, but there are some difficulties in carrying out valid fracture toughness tests because it is often difficult to generate correct starting and propagating crack profiles (Roebuck and Lord 1990). If the toughness is considered in terms of high strain rate impact, the impact energy absorbed is quite low, reflecting the susceptibility of low-ductility materials to unstable crack propagation under impact conditions.

3.5 Creep and Fatigue

Elevated-temperature creep behavior is usually considered in terms of the minimum steady-state creep rate that a material will sustain at a particular stress and temperature (see *Creep and Creep-Fatigue of Metal-Matrix Composites**). Generally, the creep resistance of particle-reinforced composites is better than that of the equivalent unreinforced material. Since the particles are elastic also at creep temperatures, creep of the matrix alloy must be the dominant process, yet there are often large discrepancies between matrix and composite for the activation energies for creep, and the stress exponent of the creep. These could indicate that very different creep mechanisms are involved, but this is probably not the case. Many of the composites have limited ductility under the creep conditions being tested, which often results in a poorly defined or nonexistent steady-state creep region. Also, it is now generally accepted that the reinforcing particles result in an increased threshold stress, below which the creep rate is essentially zero, and this accounting for the threshold stress often allows the creep behavior of the composite and the unreinforced matrix to be rationalized. It is still not clear what controls the threshold stress in particle-reinforced composites. Thermal cycling in the absence of stress does not produce a significant shape change, which is different to the situation in continuous fiber-reinforced composites. In particle-reinforced composites the residual stresses are quite low (especially at elevated temperature) and the mechanical and physical properties are isotropic, so there is no significant driving force for a shape change with thermal cycling.

An improvement in fatigue behavior is often suggested as an advantage of particle-reinforced composites, but this is an over simplistic generalization (see *Fatigue of Particle Reinforced Materials**). Stress-controlled fatigue behavior, as typified by the stress–fatigue (*S*–*N*) plot, usually indicates an improvement in fatigue life, particularly at low stresses and long lives. This reflects the higher elastic modulus of composites, but at higher stresses the fatigue lives tend to converge in response to the reduced composite ductility. Fatigue crack growth curves also show an improvement at lower stress intensity levels because the threshold stress intensity for crack growth is higher in composites. However, if fatigue behavior is examined in strain control, where the fatigue life is measured for a particular strain amplitude, composites show inferior fatigue life, particularly as the strain amplitude increases. This mainly reflects the higher modulus and work hardening rate in composites.

4. Concluding Remarks

Particle-reinforced metal matrix composites offer a useful improvement in some of the mechanical properties of metals, particularly the elastic modulus, without incurring the fabrication difficulties and cost penalties associated with continuous fiber-reinforced composites. There are often modest gains in some physical properties, such as a reduction in the coefficient of thermal expansion (see *Composites, Physical Properties of**), and improvements in frictional and wear behavior with some systems (see *Polymer Composites, Friction and Wear of**). Their major disadvantage, as with most composites, is their reduced ductility and toughness when compared with unreinforced metals. Nevertheless, particle-reinforced metal matrix composites are now finding commercial niche markets that take advantage of the property improvements they provide (see *Metal Matrix Composites: Applications**).

Bibliography

Clyne T W, Withers P J 1993 *An Introduction to Metal Matrix Composites*. Cambridge University Press, Cambridge

Cox B M, Doutre D, Enright P, Provencher R 1993 Advances in the commercialization of shape casting of Al composites. In: *Proc. 2nd Int. Conf. Cast Metal Matrix Composites*, Tuscaloosa, AL, pp. 88–105

Davidson D L 1991 In: Everett R K, Arsenault R J (eds.) *Metal Matrix Composites: Mechanisms and Properties*. Academic Press, Boston, pp. 217–34

Dixon W, Lloyd D J 1996 Wrought DURALCAN particle reinforced metal matrix composites. In: Rohatgi P K (ed.) *Processing, Properties and Applications of Cast Metal Matrix Composites*. TMS, Warrendale, PA, pp. 259–70

Halpin J C 1984 *Primer On Composite Materials: Analysis*. Technomic, Lancaster, PA

Hatch J E 1984 *Aluminum, Properties and Physical Metallurgy*. American Society for Metals, Metals Park, OH

Lloyd D J 1991 Aspects of fracture in particulate reinforced metal matrix composites. *Acta Metall. Mater*. **39**, 59–71

Lloyd D J 1995 Factors influencing the tensile ductility of melt processed particle reinforced composites. In: Lewandowski J J, Hunt Jr. W H (eds.) *Intrinsic and Extrinsic Fracture Mechanisms in Inorganic Composite Systems*. TMS, Warrendale, PA, pp. 39–47

Lloyd D J 1997 In: Mallick P K (ed.) *Composites Engineering Handbook*. Dekker, New York, pp. 631–69

Lloyd D J, Burke H, Farahbakhsh 1994 Temperature dependence of modulus in particle reinforced composites. *Mater. Sci. Technol*. **10**, 257–61

Lloyd D J, Lagacé H P, McLeod A D 1990 Interfacial phenomena in metal matrix composites. In: Ishida H (ed.) *Proc. ICCI-III—Controlled Interphases in Composite Materials*. Elsevier, London, pp. 359–76

Metcalfe A G 1974 *Composite Materials: Vol. 1. Interfaces in Metal Matrix Composites*. Academic Press, New York, pp. 65–123

Mortensen A 1990 A review of the fracture toughness of particle reinforced aluminum alloys. In: Masounave J, Hamel F G (eds.) *Fabrication of Particulates Reinforced Metal Composites*. ASM International, Materials Park, OH, pp. 235–43

Prangnell P B, Stobbs W M 1991 The effect of internal stresses on precipitation behavior in particulate reinforced Al matrix MMCs. In: Hansen N, Jensen D J, Leffers T, Lilholt H, Lorentzen T, Pedersen A S, Pedersen O B, Ralph B (eds.) *12th Risø Int. Symp. Metal Matrix Composites—Processing, Microstructure and Properties*. Roskilde National Laboratory, Roskilde, Denmark, pp. 603–10

Roebuck B, Lord J D 1990 Plane strain fracture toughness test procedures for particulate metal matrix composites. *Mater. Sci. Technol*. **6**, 1199–209

Taya M, Lulay K E, Lloyd D J 1991 Strengthening mechanisms of a particulate metal matrix composite by quenching. *Acta Metall. Mater*. **39**, 73–87

Withers P J, Stobbs W M, Pedersen O B 1989 The application of the Eshelby method of internal stress determination to short fibre metal matrix composites. *Acta Metall*. **37**, 3061–84

D. Lloyd
Alcan International Ltd., Kingston, Ontario, Canada

Metal Matrix Composites, Recycling of

Recycling of a material is defined as submitting the material to secondary processing steps that produce the same material with the same properties as the original primary product. Recycling can in some instances be conducted several times, the material undergoing, after primary processing, several cycles of secondary processing, service, and recovery without significant degradation. This is, for instance, the case for conventional aluminum alloys (Krüger 2000, ASM 1999), which can be remelted and processed with full restoration of the primary microstructure, resulting in identical material properties. Recycling can be done within the production line without any service exposure of the material, or after re-collection of components already out of service. In the latter case some degradation of the material is expected, which would have to be repaired by the recycling process. Alternatively, a material can be processed by reclamation, in which case it is processed to a different form or microstructure to produce a different, but still useful, material.

Primary processing of metal matrix composites is essentially the bonding of the composite ingredient materials, metallic matrix and reinforcement, so as to form a percolating metallic matrix containing the reinforcement as an embedded second constituent. The main issues in primary processing of metal matrix composites are to achieve bonding between the two usually nonwetting ingredient materials and to avoid deleterious interface reactions after contact between these constituent materials is established. Frequently, primary metal matrix composite processing yields near-net-shape products, at times by the very nature of the process, and also often because these composites are very difficult to machine. Depending on the composite and the component at hand, the material can be recycled, or alternatively more profitably reused in another form after reclamation.

Recycling of metal matrix composites takes advantage of the already achieved bonding between matrix and reinforcement. Recycling of metal matrix composites is attractive, given that it provides the commercial benefit of avoiding the frequently costly primary processing, and on some occasions saves expensive reinforcements. There are, however, some specific issues that must be taken into account in the recycling of these materials. Namely, it is important to avoid excessive thermal exposure in the recycling process if the constituents are mutually prone to chemical interaction. Also, because metal matrix composites are prone to accumulation of internal damage during deformation, their recycling or reuse will often require severe nondestructive testing to verify that the microstructural integrity has been preserved. Another concern that is specific to the recycling of metal matrix composites is that the reinforcement of metals is frequently only used for partial, localized, reinforcement of metal-based components. With such components, an additional effort is necessary to separate the composite from the unreinforced section, in order to allow recycling of both materials separately.

Reclamation of the composite ingredient materials requires work to separate the two constituents of the metal matrix composite. Although the interface free energy of the reinforcement in the composite often exceeds its surface free energy in air, matrix/reinforcement separation does not occur spontaneously because the complementary free surface energy of the matrix has to be formed anew in the process. This results in a positive total free energy barrier, which must be overcome in the reclamation process. This barrier can be overcome by several methods, one being the use of insoluble flux materials having a smaller surface energy with the reinforcement than the matrix. Other processes for metal matrix composite reclamation rely on gravity for separation of the composite constituents. These can be aided by blowing bubbles of an appropriate gas through the molten composite, so as to separate the reinforcement by attachment and flotation with gas bubbles (this process is akin to froth flotation processes used in ore preparation). An additional issue in reclamation of expensive reinforcements like fibers is that their properties are often degraded in the process. In some cases, when the reinforcement is far more costly than the matrix, reclamation of the reinforcement alone can be envisaged; this is achieved with relative ease by dissolving the matrix chemically.

Given the relatively wide variety of metal matrix composites and components made therefrom, recycling and reclamation of these materials raises issues that depend on the materials class. Issues specific to the two principal metal matrix composite classes are therefore more specifically addressed in what follows, together with questions specific to reclamation of composite ingredients.

1. Particle Reinforced Metals

Cast primary metal matrix composite products containing relatively low volume fractions of discontinuous reinforcements, like foundry ingots, extrusion billets, or rolling slabs, can be recycled with comparative ease. If the recovered metal matrix composite cannot be used for the same component and application at equal performance quality, reuse can be envisaged for different applications allowing a certain degree of downgrading of the performance of the material. In the case of components selectively reinforced by particulates, remelting results in dilution of the reinforcement, yielding a different composite; here again, the resulting material could be reused in a different application.

With particulate reinforcements at volume fractions less than 30%, recycling of the composite can be achieved by remelting if the constituents are not overly reactive. The foundry composite alloy AlSi9Mg/SiC/10-20p (Lloyd 1989) provides an example where in-house return rates in the production line may amount to up to 50% without property degradation (Klimowicz 1994). Some specific measures are recommended for the recycling procedure: dry preheated scrap and tools, no overheating of the melt above 750 °C, impellers to produce convection beneath the dross without formation of a vortex, and moderate fluxing with argon or SF_6 (Provencher *et al.* 1992, Chamberlain and Bruski 1998). Careful sorting and cleaning is required before the recycling of used particle reinforced metal components in a similar remelting process not yet established industrially.

Aluminum alloys of the 2xxx, 6xxx, and 7xxx series reinforced with alumina particles react at the interface even close to the solidus temperature. Therefore, any remelting process can potentially alter the composite constituent phases and properties. In the case of AA6061/Al_2O_3/10-20p it is reported (Schuster *et al.* 1993) that the amount of interface spinel formation stabilizes without significant influence on the mechanical properties of the composite. Recycling of Duralcan wrought alloys is so far only executed on a large scale by the producer (Klimowicz 1994). Composite systems such as Mg–SiC, Al–SiC with little silicon content, Al–B_4C, and Cu–SiC, are too reactive to allow remelting without severe interface reactions.

Iron-base alloys with Ti(C,N) reinforcement stabilize at the solubility limit, which depends on the concentration of other carbide-forming alloying elements.

It has been proposed to reuse particle reinforced aluminum alloys for the production of aluminum foam (Gergely *et al.* 2000). Indeed, several closed-cell aluminum foam production processes rely on the presence of ceramic particles within the melt for stabilization of the foam structure during the process (Jin *et al.* 1992, Ruch and Kirkevag 1991, Gergely and Clyne 1999). One advantage of this recycling route is that property requirements placed on composites to be used for the production of metallic foams may not be as stringent as for bulk composites.

Particle reinforced metals also have the attribute that they can be disintegrated into smaller fragments without disintegration of their microstructure. Machining chips of these materials can thus be reused to fabricate MMC components by solid state compaction methods. This solid state process has the added advantage of minimizing interfacial reactions between matrix and reinforcements.

2. *Melt-infiltrated Preforms*

Metal matrix composites with short or continuous fibrous reinforcements are usually produced by infiltration of preforms, in a process that also defines the final shape of the metal matrix composite, and which can also produce fully or partially reinforced components. Recycling of such composites would thus require preservation of the shape of the component. For example, the alumina short fiber preform ring reinforcing the crown of an aluminum alloy piston can be recuperated during remelting. Although the matrix is molten, such MMCs are mechanically stable up to about 850°C (Burkhard 1995) and can be handled. The integrity of such MMC rings can be confirmed by nondestructive testing and inserted again into the casting of a piston, which has only to achieve bonding between the composite and the matrix, as the primary casting process had already infiltrated the preform. Recycling of MMC sections of defined shape can be envisaged for alumina fiber reinforced aluminum (or magnesium alloys, provided that recuperation of the composite from the melt is rapid enough to restrict interface reactions), SiC reinforced aluminum–silicon alloys, carbon fiber reinforced magnesium, and, with some reservations, for monofilament reinforced titanium alloys.

Particle reinforced metals with particulate volume fractions greater than 40% are also frequently prepared by melt infiltration of powder preforms. It may be envisaged to reshape simple geometries of these composites to some extent by forming processes; if remelted, the volume fraction of reinforcement might be increased if some of the matrix is squeezed out during secondary pressure casting (Burkhard 1995). Alternatively, these composites could be used as a basis for the production of lower volume fraction composites by remelting and dilution by a matrix melt (Klier *et al.* 1991).

3. *Reclamation of Ingredients*

Isolated ceramic particulates or fibers can be removed from aluminum or magnesium melts by conventional salt addition or fluxing techniques as executed to remove oxide films (Provencher *et al.* 1992). By gravity settling, the fluxed ceramic ingredients float to the dross at the top of the aluminum melt, or accumulate as sediment at the bottom of a magnesium alloy. Rotary salt furnace technology is an established reclamation process to recover aluminum from various mixtures, including particle reinforced metals; however, this requires 20–50 wt.% salt (Duralcan 1990). Both wrought and foundry alloys can be recovered using this technique, even machining-chips of fiber reinforced MMC. Typically, about 80% recovery of the available aluminum can be achieved (Klimowicz 1994).

The efficiency of particle removal by fluxing is related to the probability of contact between the ceramic constituent and the flux. Duralcan, as a supplier of particulate reinforced aluminum, proposes to incorporate the salt into the melt, agitated by injecting gas (Duralcan 1996). Thus dewetting is achieved by much smaller salt additions (less than 1 wt.% for alumina and about 1.5 wt.% for SiC) and the combination with adsorption of gas to the reinforcement also accelerates particle separation by floating. This gas injection technique can also be applied to melts incorporating fiber reinforced components (Kainer 1996). Gas bubbles are adsorbed sufficiently to the fiber preforms to make these float to the surface of an aluminum melt; however, this process proves more difficult with magnesium melts. With these, dewetting can be achieved by stirring in a suitable salt; here, the preform with the incorporated salt will then sediment to the bottom of the crucible instead of floating (Kiehn *et al.* 1996). Nishida *et al.* (1999) reported on different flux systems for 10–20 vol.% alumina reinforcement in pure aluminum and in AlSiCuMg alloys (AA6061); these allow reclamation of up to 50% of the matrix. Only 20% of the AA6061 matrix could, however, be reclaimed from a 25 vol.% SiC-whisker reinforcement, presumably due to severe interface carbide formation. The preforms are partly infiltrated by the flux, causing these to float on the aluminum melt where they can be skimmed off.

Burkhard (1995) reports on centrifugal filtering of aluminum alloys reinforced by particulates, and also by short or continuous alumina fibers. Using medium-grade graphite filters in a centrifugal plasma torch furnace at accelerations above $2000\,\mathrm{m\,s^{-2}}$, more than 50% of the matrix melt can be separated from ceramic reinforcements. Further reuse can also be imagined for the highly enriched filter cakes containing up to 40 vol.% particulates or short fibers.

A melt extraction technique was reported (Lotze *et al.* 1996) to recover the aluminum melt by producing rapidly solidified needles thereof. In this method, remelted particulate or short fiber reinforced aluminum are held above the liquidus for times sufficient to segregate the heavier reinforcements to the bottom. A cooled copper disk with grooves is then dipped into the melt to extract flakes or metallic needles by fast rotation; these can then be submitted to powder compaction techniques. Again the sediment is enriched in discontinuous reinforcement and could be reused.

The reclamation of the reinforcement free of metal matrix can be achieved by dissolving the matrix chemically. The necessary effort will increase with the volume fraction of the inert reinforcement, because convection is progressively hindered. Such leached reinforcement, especially fibers, could be chopped for reuse as discontinuous reinforcements. Klimowicz (1994) mentions that from the segregates of discontinuously reinforced metals resulting from the gas injection method or from of the plasma furnace, a ceramic powder may also be recovered.

Bibliography

American Society of Materials 1999 *Proc. 4th ASM Int. Conf. Recycling of Metals*, Vienna, 17th/18th June 1999. ASM International, Metals Park, OH

Burkhard R 1995 Recycling von Metallen aus metallhaltigen industriellen Rohstoffen mittels Plasma. Ph.D. thesis, ETH-Zurich

Chamberlain B, Bruski R 1998 *Melting and Recycling of DURALCAN® MMC Material*. Duralcan-USA, San Diego, CA

Duralcan 1990 *Duralcan® Composite Casting Guidelines*. Duralcan-USA, San Diego, CA

Duralcan 1996 *Aluminium Recovery from Metal Matrix Composite Scrap*. Duralcan-USA, San Diego, CA

Gergely V, Clyne T W 1999 A novel melt-based route to aluminium foam production. In: Banhart J *et al.* (eds.) *Proc. Metal Foams and Porous Metal Structures*. Verlag MIT Publishing, Bremen, Germany, pp. 83–9

Gergely V, Degischer H P, Clyne T W 2000 Recycling of MMC and production of metallic foams. In: Clyne T W (ed.) *Comprehensive Composite Materials, Volume 3: Metal Matrix Composites*. Pergamon, Oxford, pp. 797–820

Jin I, Kenny L D *et al.* 1992 Stabilized metal foam body. *US Pat.* 5 112 697

Kainer K U 1996 Konzepte zum Recycling von Metall-Matrix-Verbundwerkstoffen. In: *Leonhardt D et al. Proc. Recycling von Verbundwerkstoffen*. Deutsche Ges.f. Materialkunde, Oberursel, Germany, pp. 39–45

Kiehn J, Oldörp F, *et al.* 1996 Wiederverwertung von Kreislauf material aus Kurzfaserverstärkten Magnesiumlegierungen. In: Leonhardt D *et al.* (eds.) *Proc. Recycling von Verbundwerkstoffen*. Deutsche Ges. f. Materialkunde, Oberursel/Germany, pp. 51–6

Klier E M, Mortensen A, Cornie J A, Flemings M C 1991 Fabrication of cast particle reinforced metals via pressure infiltration. *J. Mater. Sci.* **26**, 2519–26

Klimowicz T F 1994 The large scale commercialization of aluminum matrix composites. *J. Met.* **46** (Nov), 49–53

Krüger J 2000 *Aluminium Recycling*. Verein Deutscher Schmelzhuetten e.V., Duesseldorf, Germany

Lloyd D J 1989 The solidification microstructures of particulate reinforced Al/SiC composites. *Compos. Sci. Tech.* **35**, 159–80

Lotze G, Lehnert F, Stephani G, Kostmann C, Büttner T 1996 Rezyklieren des metallischen Anteils von Al-MMC mittels Tiegelschmelzextraktion. In: Leonhardt D, et al. (eds.) *Proc. Recycling von Verbundwerkstoffen*. Deutsche Ges. f. Materialkunde, Oberursel, Germany, pp. 45–50

Nishida Y, Izawa N, Kuramasu Y 1999 Recycling of aluminum matrix composites. *Metall. Mater. Trans.* **30A**, 839–44

Provencher R, Riverin G, *et al.* 1992 Recycling of Duralcan Aluminium Metal Matrix Composites. In: *Proc. Adv. Production and Fabrication of Light Metals and Metal Matrix Composites, Montreal*, TMS, Warrendale, PA

Ruch W W, Kirkevag B 1991 A process of manufacturing particle reinforced metal foam and product thereof. *Pat.* WO 9101387, 7.2.1991

Schuster D M, Skibo M D, *et al.* 1993 The recycling and reclamation of metal–matrix composites. *J. Met.* **45** (May), 26–30

H. P. Degischer

Technische Universität Wien, Austria

Metal Matrix Composites: Applications

Metal matrix composites (MMCs) comprise a relatively wide range of materials defined by the metal matrix, reinforcement type, and reinforcement geometry (see *Metal Matrix Composites: Matrices and Processing**). As regards the matrix, most metallic systems have been explored for use in MMCs, including aluminum, beryllium, magnesium, titanium, iron, nickel, cobalt, copper, zinc, and silver. By far the largest usage is in aluminum matrix composites.

As for the reinforcement, the materials used are typically ceramics since they provide a very desirable combination of stiffness, strength, and relatively low density. Candidate reinforcement materials include SiC, Al_2O_3, B_4C, TiC, TiB_2, graphite, and a number of other ceramics. In addition, there has been work on metallic materials as reinforcements, notably tungsten and steel fibers.

The morphology of the reinforcement material is another variable of importance in MMCs. The three major classes of reinforcement morphology are continuous fiber, chopped fiber or whisker, and particulate. Typically, the selection of the reinforcement morphology is determined by the desired property/cost combination. Generally, continuous-fiber-reinforced MMCs provide the highest properties in the direction of the fiber orientation but are the most expensive. Chopped-fiber- and whisker-reinforced materials can produce significant property improvements in the plane or direction of their orientation, at somewhat lower cost. Particulates provide a comparatively more moderate but isotropic increase in properties and are typically available at the lowest cost.

By adding to the three variables of metallic matrix, reinforcement material, and reinforcement morphology the further parameters of reinforcement volume fraction, orientation, and matrix alloy composition and heat treatment, it is apparent that there is a very wide range of available material combinations and resultant properties. The specifics of MMC systems, their manufacturing processes, and properties are covered elsewhere; this article discusses how MMCs have been, and are being, used in engineering, with a focus on principal application areas.

In the vast majority of these applications, MMC materials have been implemented after displacing other materials, such as unreinforced metals and alloys or polymer matrix composites. Reasons for this substitution vary widely, and include increases in modulus, strength, wear resistance or damping capacity, as well as reductions in thermal expansivity or density, and the cost competitiveness of MMCs.

1. Applications

Applications of MMCs can be conveniently grouped by market area, since each area requires sets of properties and costs that result in particular subsets of MMC materials being of primary interest. The primary markets for MMCs are divided into aerospace, automotive, commercial and industrial products, and electronic packaging and thermal management; all these represent the areas of greatest interest. In each of the sections that follow, a summary of applications either in production or in an advanced prototype stage will be covered to provide a realistic perspective on the status and potential of these materials in the selected market areas. Further information can be obtained in the specific references in each section as well as reviews on the subject (Zweben 1992, Koczak *et al.* 1993, Harrigan 1994, Eliasson and Sandstrom 1995, Peel 1995, Kevorkijan 1998, Maruyama 1999, Froes 1999).

1.1 Aerospace

Aerospace applications were the original driving force for MMC development. This was due to the quest for weight reduction for improved performance and payload capability combined with the high value placed on weight savings.

(*a*) *Aircraft structure.* A significant portion of aircraft structure is designed by stiffness, so as a result many of the aircraft structure applications for MMCs seek to capitalize on the increased specific stiffness that these materials provide. Aircraft structural applications actually require a combination of properties, including adequate strength, damage tolerance (including ductility, fracture toughness, and fatigue resistance), and corrosion resistance. MMCs typically have lower damage tolerance properties than their unreinforced matrix metal counterparts, and hence the extent of structural application in primary structures has been limited. More extensive application depends on the development of cost-effective large-scale MMC production processes that result in materials with improved combinations of properties over currently available materials. There have been selected aircraft structure applications that have been realized. These are exclusively aluminum-based MMCs, and will be discussed briefly below.

SiC particulate-reinforced aluminum MMCs are used in the Lockheed F-16 fighter aircraft. One application is in the ventral fins of the aircraft, which provide lateral stability during high angle of attack maneuvers. A 6092/SiC/17.5p MMC sheet material replaced the unreinforced aluminum skins in the honeycomb structure of the fin and, due to the increased specific stiffness of the MMC material, increased the service life significantly. The same material is used in the fuel access door covers for the F-16. Unreinforced aluminum access covers were experiencing cracking due to overload, and again the increased specific stiffness property of the MMC allowed this material to be used successfully (Maruyama 1999).

The Eurocopter rotor sleeve is another application that has been reported (Hurley 1995). A forged SiC particulate-reinforced 2009 alloy having good stiffness and damage tolerance has been used. It is a replacement for a titanium part with a reduction in weight and production cost.

Replacement of a carbon-fiber-reinforced polymer tube by a SiC particulate-reinforced aluminum extruded tube for a floor support strut in Airbus aircraft has also been reported (Hurley 1995, Peel 1995). The driving force is improved damage tolerance and reduced cost.

(*b*) *Aeroengines.* Gas turbine engines also have substantial needs for lightweight, high-performance materials. This application area has a rich history of innovative materials developments, and the expectation that MMCs would be an important materials class in aeroengines is not surprising. In the case of this application, both specific stiffness and specific strength, especially at elevated temperature, are the critical properties. But similarly to the case for aircraft structure, other properties, including fatigue, creep, and oxidation resistance, are also very important.

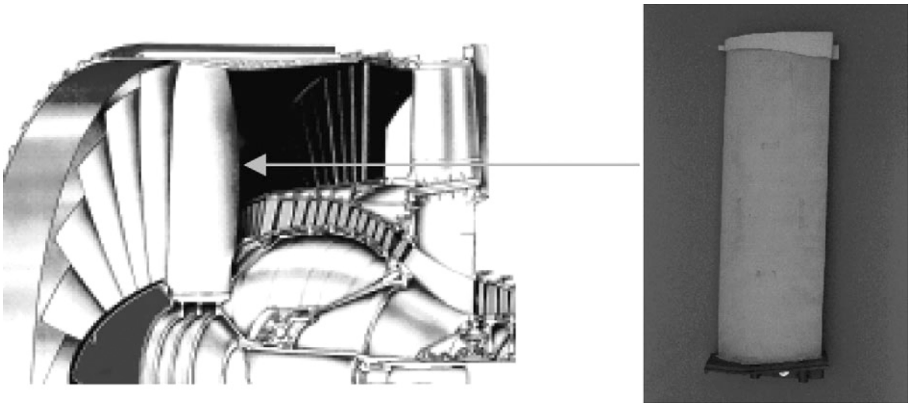


Figure 1
Fan exit guide vane for Pratt and Whitney PW4000 engine made from extruded 6092/SiC/17.5p material.

Aluminum MMCs are being applied in aeroengines, specifically in the area of the fan exit guide vanes for the Pratt and Whitney PW4000 series engines that power many of the Boeing 777 commercial aircraft (Fig. 1). In this application, a 6092/SiC/17.5p-T6 extruded material replaced a graphite/epoxy material in order to achieve improved resistance to foreign object damage and erosion. Significant acquisition and lifecycle cost savings have been reported (Maruyama 1999).

Titanium MMCs have also been explored for replacement of nickel-based alloys (Wei 1992). While there was extensive development effort in this area, actual application has not been realized.

(*c*) *Space structures.* The premium for weight savings in structures that will be launched into space is very high. Typically, space structures can be viewed as platforms for electronic components, and hence specific stiffness and, to a lesser degree, specific strength are important. In addition, the physical properties of the materials, particularly coefficient of thermal expansion and thermal conductivity, are important due to the thermal excursions that occur in orbit and the need to maintain dimensional accuracy.

An early application for MMCs was in the Space Shuttle. In this case, continuous boron-fiber-reinforced aluminum is used in struts comprising the fuselage frame (Kinner 1980).

On the Hubble Space Telescope, a high-gain antenna boom constructed of continuous graphite-fiber-reinforced aluminum is used. The primary driving force for this material was the dimensional stability that can be engineered into the composite through the addition of graphite fibers with a negative coefficient of expansion in the fiber direction at an appropriate volume fraction to balance the positive coefficient of expansion of the aluminum matrix. This structure replaced an aluminum waveguide and graphite/epoxy tube assembly with a single part with a weight saving of 30% (Harrigan 1994).

(*d*) *Other aerospace applications*. SiC particulate-reinforced aluminum was used in the form of extruded sections for electronic racks. Specifically, a 6061/SiC/25p material replaced unreinforced aluminum racks in a military aircraft with a 30% weight saving (Harrigan 1994).

In addition, the ability to tailor the coefficient of thermal expansion of MMCs has been useful in replacing materials in structures requiring thermal stability. General Electric Co. designed an advanced composite optical system gimbal, using 6061/SiC/40p MMC in selected areas requiring low expansion and high wear resistance (Harrigan 1994).

Another application using the ability to match coefficient of thermal expansion with mating materials was the use of 6061/SiC/40p MMC in the covers for an inertial guidance unit. In this application, the MMC replaced beryllium that had to be machined from a solid block. The MMC material could be forged to near-net shape, with final machining only required. In addition, the machining of the MMC could be done in conventional machine shops, since it does not have the toxicity concerns of beryllium machining (Harrigan 1994).

1.2 Automotive

In the automotive market, properties of interest to the automotive engineer include the increased specific stiffness, wear resistance, and improved high-cycle fatigue resistance (Allison and Cole 1993). While weight savings are also important in automotive applications, the need for achieving performance improvements with much lower cost premiums than tolerated by aerospace applications drives attention toward low-cost materials and processes. There has seen successful application of MMCs in selected automotive applications in which the combination of properties and cost satisfied particular needs.

(*a*) *Engine*. Replacement for steel and cast iron in engine applications relies on the increased specific stiffness, improved wear resistance, and, in some cases, on the increased high-cycle fatigue resistance provided by MMCs.

A watershed application for aluminum MMCs was the Toyota piston for diesel engines (Donomoto, *et al.* 1983). The part consists of selective reinforcement of the aluminum alloy by a chopped fiber preform in the ring-groove area in order to provide improved wear and thermal fatigue resistance. These pistons have been in commercial production in Japan since the early 1980s.

Other piston applications have included the use of SiC particulate-reinforced aluminum forgings in racing applications. Due to the lower coefficient of thermal expansion of the MMC, reduced clearances between the piston and cylinder wall are possible. Based on trials of MMC pistons in drag-racing bikes, improved performance compared to conventional hypereutectic Al–Si alloys results (Harrigan 1994).

Another application in which selective reinforcement is employed is in the Honda Prelude 2.3-liter engine (Hamajima *et al.* 1990). In this case hybrid preforms consisting of carbon and alumina fibers were infiltrated by molten aluminum to form the cylinder liners during the medium pressure squeeze casting process for the engine block.

Other drivetrain components, particularly connecting rods, have been a focus of development (Harrigan 1994). By reducing the mass of the connecting rod/piston assembly, the objectionable secondary shaking forces that can develop, particularly in smaller engines, can be reduced. In addition, lower reciprocating loads should lead to lower loads on the crankshaft, lower friction losses, and increased fuel economy or performance can be realized (Allison and Cole 1993). No commercial applications of connecting rods in high-volume vehicles have been achieved, largely because of the difficulty in obtaining a material with the necessary high-cycle fatigue performance and low cost combination. While prototype connecting rods from hot forged aluminum MMC have been prototyped and tested, further cost reduction is required.

(*b*) *Brake system*. Aluminum-based MMCs offer a very useful combination of properties for brake system applications to replace cast iron. Specifically, the wear resistance and high thermal conductivity of aluminum MMCs enable substitution in disk brake rotors and brake drums with attendant weight savings of the order of 50–60%. Since the weight reduction is in unsprung weight, it also reduces inertial forces and provides additional benefit. In addition, lightweight MMC rotors provide increased acceleration and reduced braking distance. It is reported that, based on brake dynamometer testing, MMC rotors reduce brake noise, do not wear as much, and have more uniform friction over the entire testing sequence compared to cast iron rotors (Allison and Cole 1993). Commercial applications of aluminum MMC in brake rotors and drums have been achieved; these include the rear drums on the GM EV-1 as well as the rotors on the Plymouth Prowler and Lotus Elise. Examples of these components are shown in Fig. 2.

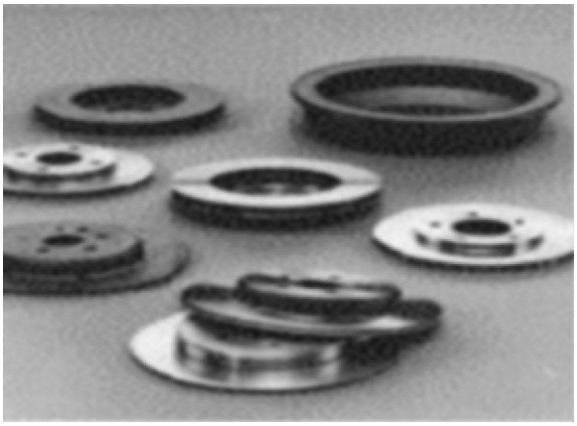


Figure 2
Brake rotor and brake drum components from cast SiC particulate-reinforced aluminum MMC.

Application of aluminum MMCs in automotive racing applications, where higher priced material is acceptable for improved performance, has also been reported. Brake calipers for Formula 1 racing cars produced from a 2124/SiC/25p MMC provide less displacement, more leverage, and quicker stops due to the increased material stiffness (Hurley 1995).

(*c*) *Driveshaft*. Use of aluminum MMCs in the driveshaft takes advantage of the increased specific stiffness of these materials. Conventional driveshafts, whether aluminum or steel, are constrained by the speed at which the shaft becomes dynamically unstable. The critical speed of the driveshaft is a function of the length, inner and outer radius, and specific stiffness. In vehicles with packaging constraints, that do not allow increased driveshaft diameter, MMCs offer a desirable solution. Use of MMCs enables longer driveshaft lengths at a given diameter or smaller diameter shafts at a given length. As a result of these benefits, driveshafts comprised of 6061/Al_2O_3 materials produced by stir casting and subsequent extrusion into tubes have been applied in the GM S-T truck and Chevrolet Corvette.

(*d*) *Other automotive applications*. MMCs, specifically those based on aluminum matrices, are candidates for use in brake calipers, pump housings, gears, valves, brackets, pulleys, turbocharger and supercharger compressors, and suspension components (Allison and Cole 1993). In addition, they have been cited for use in clutch parts, suspension pushrods and rockers, as well as other gearbox and engine parts (Hurley 1995).

An engine cradle for the General Motors EV-1 has been prototyped in B_4C-reinforced aluminum produced by die casting (Froes 1999).

Snow tire studs have been manufactured from 6061/Al_2O_3 drawn wire. In Finland, where steel tire stud jackets have been outlawed and unreinforced aluminum has insufficient wear resistance, MMC tire stud jackets have been successfully applied (Nussbaum 1997).

1.3 Electronics Industry

The ability to tailor the coefficient of thermal expansion of MMCs through additions of ceramic reinforcement, while maintaining the thermal conductivity of the metallic base alloy, has opened up an entirely new area of application, namely in thermal management and electronic packaging applications (Zweben 1992). MMCs produced by both powder metallurgy and liquid metal infiltration methods have been employed in thermal management applications, and are commonly known in the market as "AlSiC" materials; Fig. 3 shows an example of an electronic substrate produced from SiC particle-reinforced aluminum.

An emerging application in which MMCs would appear to offer advantages is in substrates for hard disk drives. These substrates are predominantly made from aluminum, with some selected applications using glass or ceramic materials. As the disk drive developers seek to spin the hard disks at higher speeds for faster data transfer, the upper speed is limited by the occurrence of disk flutter, which is determined by material modulus. Increasing the elastic modulus of the substrate could allow higher speeds without flutter, and aluminum MMCs have been prototyped for this application (Froes 1999).

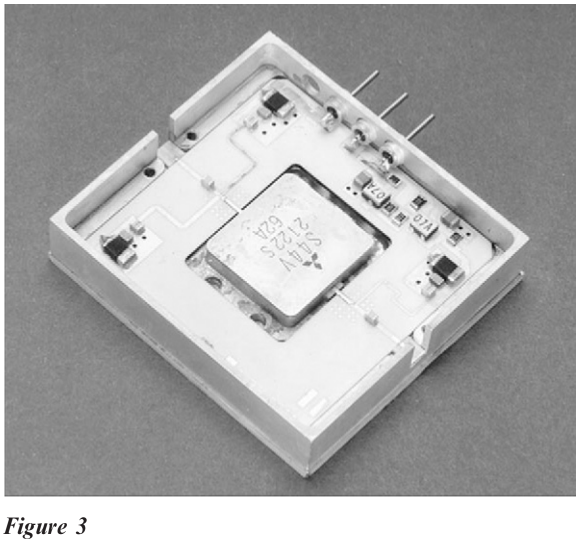


Figure 3
Electronic package manufactured from AlSiC MMC.

1.4 Other Commercial and Industrial Products

In the recreational products market, improved performance (real or perceived) is highly valued. As a result, many of the materials that gain favor in the aerospace market are also applied in recreational products. The primary property of MMCs that is of interest for this market area is the increased specific stiffness.

A range of materials has been targeted at the golf equipment market. Both shafts and club heads for both woods and irons are of interest, taking advantage of the high specific stiffness and strength properties of MMCs (Hurley 1995, Froes' 1999).

Applications in the bicycle market have included frames, front forks, handlebars, and seat posts (Hurley 1995, Peel 1995). In this market the stiffness and durability of MMCs are desired, and for one 6061/Al_2O_3 extruded tube product, it has been reported that over 50 000 aluminum MMC frames per year were sold by one bicycle company (Nussbaum 1997).

One application of SiC particle-reinforced aluminum that was developed but never applied was a mast for the America's Cup racing yacht *Stars and Stripes '88*. The mast was made from a 6061/SiC/20p MMC that was to replace a thicker walled unreinforced aluminum mast with the attendant weight savings (Harrigan 1994).

B_4C-reinforced aluminum has promise in nuclear shielding applications. The B-10 isotope present in B_4C naturally absorbs neutron radiation, and as a result this type of MMC is being considered for use in storage casks that will contain spent fuel rods from nuclear reactors (Froes 1999).

Use of aluminum MMCs in an industrial application requiring a lower weight material for improved precision has been reported. Specifically, a 6 m long MMC needle was used in a carpet-weaving machine to replace a steel needle (Hurley 1995).

2. Concluding Remarks

MMCs have found selected applications in areas that can cost-effectively capitalize on improvements in specific stiffness, specific strength, fatigue resistance, wear resistance, and coefficient of thermal expansion. New inroads will depend not only on further enhancement of the properties of these materials but also on further reductions in raw materials as well as part-manufacturing costs.

Bibliography

Allison J E, Cole G S 1993 Metal matrix composites in the automotive industry. *JOM* **45**, 19–24

Donomoto T *et al*. 1983 SAE Paper 830252, SAE

Eliasson J, Sandstrom R 1995 *Key Engineering Materials*. **104–7**, 3–36

Froes F H 1999 *Light Metal Age* **57**(3–4), 48–61

Hamajima K, Tanaka A, Suganama T 1990 *JSAE* **11**, 80–4

Harrigan W C 1994 In: Ochiai S (ed.) *Handbook of Metallic Composites*. Marcel Dekker, New York, pp. 759–73

Hurley S 1995 MMCs for automotive applications. *MBM* **XX**, 54–5

Kevorkijan M V 1998 Metal matrix composites. *Am. Ceram. Soc. Bull*. Dec., 53–9

Kinner W K 1980 *Mater. Eng*. **91**, 64–6

Koczak M J, *et al*. 1993 In: Suresh S, et al. (ed.) *Fundamentals of Metal Matrix Composites*. Butterworth-Heinemann, Boston, pp. 297–326

Maruyama B 1999 Progress and promise of aluminium composites. *Adv. Mater. Processes* **156** (1), 47–50

Nussbaum A I 1997 *Light Metal Age* **55** (1–2), 1–2

Peel C J 1995 In: Lee E W, Kim N J, Jata K V, Bader M G (eds.) *Light Weight Alloys for Aerospace Applications III*. TMS, Warrendale, PA, pp. 191–205

Wei W 1992 *Metals Mater*. **8** (8), 430–5

Zweben C 1992 Metal–matrix composites for electronic packaging. *JOM* **44**, 15–23

W. H. Hunt
Aluminum Consultants Group Inc., Murrysville, Pennsylvania, USA

Metal Matrix Composites: Matrices and Processing

The term metal matrix composite (MMC) covers various types of system, and also a wide range of scales and microstructures. Common to them all is a metallic matrix, which is normally contiguous. The reinforcing constituent is in most cases a ceramic, although there are exceptions to this and MMCs can be taken to include materials "reinforced" with relatively soft and/or compliant phases, such as graphitic flakes, lead particles, or even gases. It is also possible to use refractory metals, intermetallics, or semiconductors rather than true ceramics. MMC types are commonly subdivided according to whether the reinforcement is in the form of (i) particles, which are at least approximately equiaxed; (ii) short fibers (with or without a degree of alignment); or (iii) long aligned fibers. Choice of matrix and reinforcement, specification of the way in which the composite material is to be synthesized, and the manner in which a stock item or component is to be fabricated from this material are key interwoven issues for technologists interested in product development. In this article, a brief overview is given of the state of the art in these areas.

1. *MMC Systems and Issues for their Processing*

A flow chart of the routes used for MMC processing is given in Fig. 1. The individual composite production operations are briefly described in Sect. 2. A summary is given of the different types of MMC, the matrices and reinforcements primarily involved, and associated issues related to processing—particularly the specifics of chemical compatibility.

1.1 Particulate MMCs

Particulate MMCs are being used or developed for a range of industrial applications. While these are often focused on aluminum alloy matrices, titanium-, iron-, and magnesium-based systems are also of interest. The particulate is most commonly SiC or Al_2O_3, but others (TiB_2, B_4C, SiO_2, TiC, WC, BN, ZrO_2, tungsten, etc.) have been investigated. Chemical reaction during processing can occur in some cases. Silicon carbide can be particularly problematic in aluminum- and titanium-based MMCs. The reaction between SiC and aluminum melts has been investigated in detail and SiC reacts with titanium even during solid-state consolidation. Alumina is less reactive than SiC in aluminum alloys, but it does react quite strongly with titanium.

Magnesium is rather different from aluminum and titanium, in that it does not readily form a stable carbide. However, its affinity for oxygen is very strong. The greater stability of Al_2O_3, compared with SiC, in aluminum is therefore reversed for magnesium matrices. In general, while coatings or other surface treatments may be worth considering for fibers (particularly monofilaments), economic and practical considerations mean that particulate reinforcement is normally introduced into MMCs in the virgin state. This may, however, be such that a surface oxide layer is present and deliberate thickening of this layer, for example by heat treatment in air, has in some instances been found to have a beneficial effect on interfacial bonding or other characteristics.

Particulate MMCs are most commonly manufactured either by a melt incorporation and casting technique or by powder blending and consolidation. Other routes include reactive processing or spray codeposition. Quality control objectives include the elimination of excessive interfacial reaction during processing, particularly for melt routes, and also the avoidance of microstructural defects such as poor

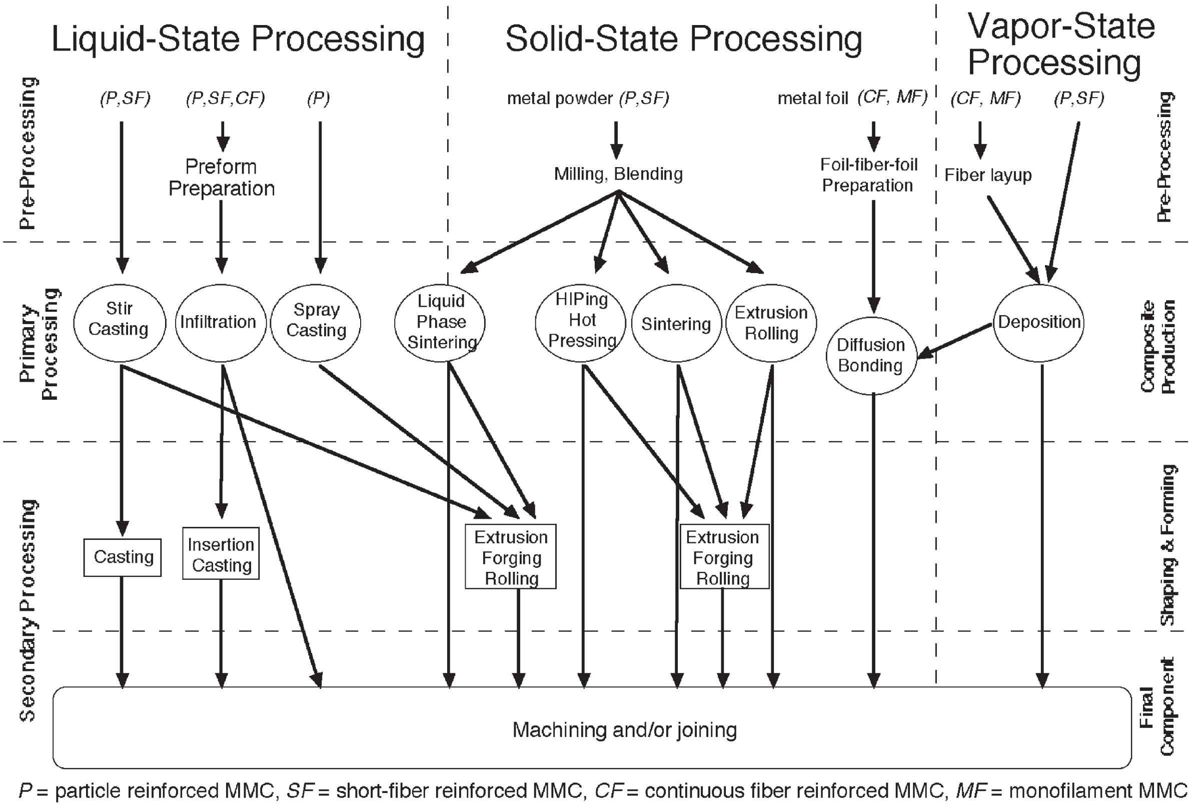

Liquid-State Processing
Solid-State Processing
Vapor-State Processing
Pre-Processing
Primary Processing
Secondary Processing
Pre-Processing
Composite Production
Shaping & Forming
Final Component
(P,SF)
(P,SF,CF)
(P)
Preform Preparation
metal powder (P,SF)
Milling, Blending
metal foil (CF, MF)
Foil-fiber-foil Preparation
(CF, MF)
(P,SF)
Fiber layup
Stir Casting
Infiltration
Spray Casting
Liquid Phase Sintering
HIPing Hot Pressing
Sintering
Extrusion Rolling
Diffusion Bonding
Deposition
Casting
Insertion Casting
Extrusion Forging Rolling
Extrusion Forging Rolling
Machining and/or joining
P = particle reinforced MMC, SF = short-fiber reinforced MMC, CF = continuous fiber reinforced MMC, MF = monofilament MMC

Figure 1
Overview flow chart of MMC processing routes (courtesy of A. Mortensen, C. San Marchi, and H. P. Degischer).

interfacial bonding, internal voids, and clustering of the reinforcement. Typically, for structural applications reinforcement particles are about 10–20 μm in diameter and constitute about 10–30% by volume of the material, although MMCs in which the values concerned lie outside of these ranges have been studied and are available commercially (particularly finer particles and higher particle contents). For electronic substrate applications, far higher particle contents, around 70% by volume, are commonly employed. These are obtained using multimodal particle size blends.

1.2 Short Fiber MMCs

Short fiber MMCs first attracted widespread attention in the mid-1980s, with the development of aluminum diesel engine pistons selectively reinforced with short alumina (e.g., ICI's "Saffil") fibers. Other rather similar (alumino-silicate) fibers have also been employed for applications of this type. These fibers have a fine-grained polycrystalline microstructure. Typical fiber diameters are a few micrometers and they are initially produced in lengths of several hundred micrometers. Components are commonly produced by melt infiltration. Interfacial characteristics are partly dependent on the degree of reaction during processing, which is in turn affected by the surface chemistry of the fibers. For example, Saffil fibers have a thin silica-rich surface layer that tends to react with an aluminum melt during processing, particularly if magnesium is present. Some secondary processing, such as forging and extrusion, can be carried out under appropriate conditions. However, their formability is in general markedly inferior to that of particulate MMCs. Certain property advantages over particulate MMCs are common, however, particularly in terms of resistance to creep and, to a lesser extent, wear.

Superior mechanical properties are obtained if the fine-grained structure can be replaced by that of a single crystal. The so-called "whiskers" that result when fine, slender fibers are produced as monocrystals created considerable interest as early as the 1960s. Whiskers are usually ~1 μm in diameter, with aspect ratios of up to several hundred. Their tensile strengths are often very high. While they were originally expensive to manufacture, relatively cheap methods for the production of SiC whiskers have been developed. However, while excellent properties have been reported for whisker-reinforced MMCs, work in this area decreased in the early 1990s and there has been little commercial exploitation. This is largely a consequence of handling difficulties. These commonly arise with very fine fibers, which tend to form tenacious ball-like structures and are difficult to orient in a controlled way. However, the most significant problem of this type concerns perceived health hazards. Whiskers and whisker fragments in the slightly submicrometer size range can become airborne very readily and are likely to reach and damage the lungs.

1.3 Long Fiber MMCs

Several long fiber MMCs systems have been investigated and some have been used in certain applications. However, as a consequence of processing difficulties and of constraints on ductility and toughness, this usage has in general remained limited. The term "multifilament" refers to relatively small-diameter (~5–30 μm) fibers, which are flexible enough to be handled as tows or bundles. These can be woven, braided, filament wound, etc. The materials concerned include carbon, SiC, and various oxides. Other multifilaments are in common use, but most are either unable to survive the elevated temperatures involved in MMC production (e.g., polymeric and organic fibers) or are of limited interest for MMCs because of relatively poor mechanical properties such as stiffness or creep resistance (e.g., glass fibers). Multifilament MMCs can be produced by melt infiltration, although problems can arise with unidirectionally aligned fibers if applied melt pressure transverse to the fiber axis brings them into close contact so that the melt is unable to penetrate. This problem can be reduced by introducing particulate or transverse fibers.

Carbon fibers are not popular as MMC reinforcement, primarily due to interfacial reactions and galvanic corrosion effects. Chemical reaction problems are severe for aluminum, titanium and iron alloys. With aluminum, the interfacial reaction product (Al_3C_4) is hygroscopic, so that graphite fiber-reinforced aluminum tends to undergo rapid corrosion in aqueous environments. Attempts have been made to protect carbon fibers with a surface coating, such as titanium nitride, but in general this is difficult and expensive for multifilaments. Magnesium does not form a carbide so magnesium-based carbon multifilament MMCs have received some attention. There has also been interest in carbon fiber-reinforced zinc alloys, another system in which interfacial reaction is limited. While SiC has been successful in particulate MMCs, multifilament SiC fibers suitable for incorporation into metallic matrices are not commercially available. Products on the market, under trade names such as Nicalon, tend to have high levels of free carbon and silica, leading to excessive reaction with most metallic matrices during processing. There are, however, oxide multifilaments that are fairly resistant to attack by molten metals. Notable among these are polycrystalline alumina fibers, usually composed primarily of the stable α phase, which have been studied for use in aluminum and, to a lesser extent, titanium matrices. Industrial interest in all such systems has remained low, although there has been some activity concerning reinforcement of intermetallic matrices with oxide multifilaments.

Monofilaments are large-diameter (~100–150 μm) fibers, usually produced by chemical vapor deposition (CVD) of either SiC or boron onto a core of carbon fiber or tungsten wire. Monofilaments are much less flexible than multifilaments, so they are handled as single fibers, rather than bundles, and precautions are necessary to avoid causing damage by the imposition of sharp curvature during handling. Interfacial reaction, however, is less of a concern, since the proportion of the fiber consumed is much less than for multifilaments. There is also greater scope for tailoring the surface chemistry and introducing surface coatings, which can be done as part of the fiber production process. For example, coatings have been developed for SiC monofilaments to be incorporated into titanium matrices. Thick graphitic coatings (which are progressively consumed but prevent defects from forming on the fiber itself) have been popular and there has also been work on various duplex layers, such as TiB_2/C and Y/Y_2O_3, which are designed to slow interfacial reaction rates down to very low levels.

1.4 Cermets

Cermets, which have microstructures comprising an assembly of ceramic particles bonded together by a small proportion of a metallic phase, can be regarded as a special type of MMC. A key to the success of cermets in industrial applications concerns their ease of processing. They are commonly produced by blending of ceramic and metallic powders followed by liquid-phase sintering. Typically, the ceramic particles are 1–10 μm in diameter. Blending involves a milling operation, which tends to coat the ceramic particles with metal. This is usually followed by cold isostatic pressing or injection molding to give the required shape, and then holding at a suitable temperature under vacuum, inert gas, or hydrogen. During liquid-phase sintering, particle rearrangement occurs, driven by capillarity forces. These may effect densification, depending on the degree of wetting. In some cases, notably for the oxide-based cermets, it is often necessary to impose unixial or hydrostatic pressure in order to eliminate porosity. This is reminiscent of melt infiltration processes used to form MMCs, but the very high ceramic content and fine scale of the structure means that in the absence of wetting very high pressure might be necessary to ensure that liquid flows into all the cavities.

One approach to the problem of inadequate wetting is to deliberately promote selected chemical reactions during sintering. For example, cermets such as TiAl–TiB_2 can be produced via the "XD" process by reacting powder blends such as aluminum, titanium, and boron. More details of reactive processing techniques are given in Sect.2.1(d). However, pronounced chemical reaction can cause difficulties in the form of uncontrolled volume changes, heat release, and undesirable phase formation. Fortunately, the carbides and nitrides concerned (e.g., WC, TiC, etc.) are in general more stable than the carbides or nitrides of the binder metals (e.g., cobalt), so there are few problems of pronounced chemical attack during sintering.

Some metal ion exchange may occur, and in certain cases there is a degree of dissolution of the ceramic in the metal and reprecipitation on solidification. Such processes probably have the effect of raising the bond strength, but for most systems there is little danger of thick, brittle interfacial reaction zones forming during sintering. Of course, this is largely because cermet systems have been chosen partly with this chemical compatibility in mind, since there are not the same constraints of economics and matrix mechanical properties that operate when choosing the constituents for MMCs. The aluminum-, titanium-, and magnesium-based matrices commonly used for MMCs are attractive in these terms, but are very reactive in the liquid state, which often leads to problems during this type of processing.

1.5 Metallic Foams

Metallic foams have attracted considerable interest—partly because processing advances have led to material becoming available at competitive cost and partly because attractive property combinations can be obtained, particularly in terms of specific stiffness and specific energy absorption. Several approaches are available for the production of metallic foams. Some of these, particularly those generating closed-cell structures, involve processing the metal in the liquid or semisolid state. A problem then arises from the low viscosity of liquid metals. Depending on the processing methodology employed, it is often necessary to stabilize the foam against cell coarsening and collapse while the metal is in a liquid or semiliquid state. This is commonly achieved by introducing a dispersion of ceramic into the melt, either as oxide films or as ceramic particles, which increases its viscosity. MMC feedstock can thus be used for melt route production of metallic foams (see Sect. 2.1(d)).

2. Processing Routes

As depicted in Fig. 1, processes can be classified according to whether the matrix is in the liquid, solid, or vapor phase while it is being combined with the reinforcement. The individual composite production operations are briefly outlined below under these groupings.

2.1 Liquid-state Processing

(*a*) *Stir casting*. This involves stirring the melt with solid ceramic particles and then allowing the mixture to solidify. This can usually be done using fairly conventional processing equipment and can be carried out on a continuous or semicontinuous basis. A concern is to ensure that good particle *wetting* occurs. Difficulties can arise from the increase in viscosity on adding particles or, especially, fibers to a melt. However, this increase is typically only by a factor of two or so with up to about 20 vol.% particulate, provided the particles remain well dispersed. This viscosity is sufficiently low to allow casting operations to be carried out. Microstructural inhomogeneities can arise, notably particle agglomeration and sedimentation in the melt. Redistribution as a result of *particle pushing* by an advancing solidification front can also be a problem. This is reduced when solidification is rapid, both as a result of a refinement in the scale of the structure and because there is a critical growth velocity, above which solid particles should be enveloped rather than pushed.

Stir casting usually involves prolonged liquid–ceramic contact, which can cause substantial interfacial reaction. This has been studied in detail for Al–SiC, in which the formation of Al_4C_3 and silicon can be extensive. This both degrades the final properties of the composite and raises the viscosity of the slurry, making subsequent casting difficult. The rate of reaction is reduced, and can become zero, if the melt is silicon rich, either by prior alloying or as a result of the reaction. The reaction kinetics and silicon levels needed to eliminate it are such that it has been concluded that casting of Al–SiC_p involving prolonged melt holding operations is suited to conventional (high-silicon) casting alloys, but not to most wrought alloys.

(*b*) *Squeeze infiltration*. Liquid metal is injected into the interstices of an assembly of short fibers, usually called a *preform*. Commonly, the preform is designed with a specific shape to form an integral part of a finished product in the as-cast form. Preforms are commonly fabricated by sedimentation of short fibers from liquid suspension. The process can also be adapted for production of particulate MMCs. In order for the preform to retain its integrity and shape, it is often necessary for a *binder* to be used. Various silica- and alumina-based mixtures are popular as high-temperature binders. The binding agent is normally introduced via the suspension liquid, so that it deposits or precipitates out on the fibers, often forming preferentially at fiber contact points, where it serves to lock the fiber array into a strong network.

The pressure required for infiltration can readily be calculated on the basis of the necessary meniscus curvature and corrections can be made for melt/fiber wetting. In practice, substantial pressures in the megapascal region are likely to be needed. In most cases, fibers do not act as preferential crystal nucleation sites during melt solidification. One consequence of this is that the last liquid to freeze, which is normally solute-enriched, tends to be located around the fibers. Such prolonged fiber/melt contact, often under high hydrostatic pressure and with solute enrichment, tends to favor formation of a strong interfacial bond. Other forms of defect are, however, common in squeeze infiltrated composites. These include porosity and local variations in fiber content and in average alloy composition.

(*c*) *Spray deposition*. Spray deposition techniques fall into two distinct classes, depending on whether the droplet stream is produced from a molten bath or by continuous feeding of cold metal into a zone of rapid heat injection. The process was developed for building up bulk metallic material by directing an atomized stream of droplets onto a substrate. Adaptation to particulate MMC production by injection of ceramic powder into the spray has been extensively explored, although with limited commercial success. Droplet velocities typically average about 20–40 $\mathrm{m\,s^{-1}}$. A thin layer of liquid, or semisolid, is often present on the top of the ingot as it forms. MMC material produced in this way often exhibits inhomogeneous distributions of ceramic particles. Ceramic-rich layers approximately normal to the overall growth direction are often seen. This may be the result of hydrodynamic instabilities in the powder injection and flight patterns or possibly to the repeated pushing of particles by the advancing solidification front in the liquid or semisolid layer, until the ceramic content is too high for this to continue. Porosity in the as-sprayed state is typically about 5–10%.

Thermal spraying differs in several respects from melt atomization processes. Deposition rates (usually $\sim 1\,\mathrm{g\,s^{-1}}$) are slower, but particle velocities ($\sim$50–400 $\mathrm{m\,s^{-1}}$) are higher. Quenching rates for each individual splat can be very high ($\sim 10^6\,\mathrm{K\,s^{-1}}$). Porosity levels are typically at least a few percent. Thermal spraying onto arrays of fibers to form MMCs has received some attention. An attraction here is the possibility of producing composite material in an operation involving only very brief exposure to high temperatures. Provided the void content and distribution are such that full consolidation can be effected with little further heat treatment, this would allow problems of excessive fiber/matrix chemical reaction during processing to be avoided—a particularly important objective for titanium-based MMCs. Unfortunately, it has proved very difficult to spray onto fiber arrays so as to produce MMCs with acceptably low void contents and there are also problems in maintaining a uniform fiber distribution.

(*d*) *Reactive processing*. Several processes have been developed in which constituents are brought together under conditions such that a chemical reaction occurs while the mixture consolidates. In some such processes, liquid metal is introduced and progressively oxidizes. For example, the *directional oxidation* of aluminum is exploited in several processes patented under the "Dimox" trade name. Other processes consist of blending several elemental powders, and reacting these by heating with transient formation of a liquid phase; one important class of such processes are the "XD" processes. An example of such a process is to blend powders of aluminum, titanium, and boron to produce $\mathrm{Al–TiB_2}$ composites.

By making a suitable powder compact, into which liquid metal is infiltrated as it reacts with the preform, near net shape forming is possible. The product often contains residual unreacted metal, so that it is effectively an MMC. Such materials often bear a marked resemblance to cermets (see Sect. 1.4), but a distinction can be drawn in terms of certain differences in processing conditions and in the materials used. The work originated in attempts to facilitate melt infiltration, particularly with aluminum-based melts, by reducing the pressure required for infiltration. Ideally, "wetting" and/or chemical reactions occurring at the infiltration front would be such as to promote spontaneous infiltration without the need for external application of pressure. While it has proved difficult to promote rapid infiltration under these conditions, MMC products can be made in this way with good near net shape characteristics, particularly when the ceramic content is high. For example, wetting in aluminum melts can be promoted by the introduction of magnesium into the melt and nitrogen into the surrounding atmosphere (preferably in the absence of oxygen). While the thermodynamic and kinetic details are complex, reactions take place so as to promote spontaneous penetration of an aluminum-based melt into an array of ceramic particles or fibers. A schematic depiction of the process is shown in Fig. 2(a). Terms such as the "PRIMEX" process have been applied to this and related procedures. Note the bimodal size distribution of the SiC particles shown in Fig. 2(b), which facilitates the generation of high ceramic contents in the initial powder compact; this type of microstructure is typical of composites intended for applications as electronic substrates, as mentioned in Sect. 1.1.

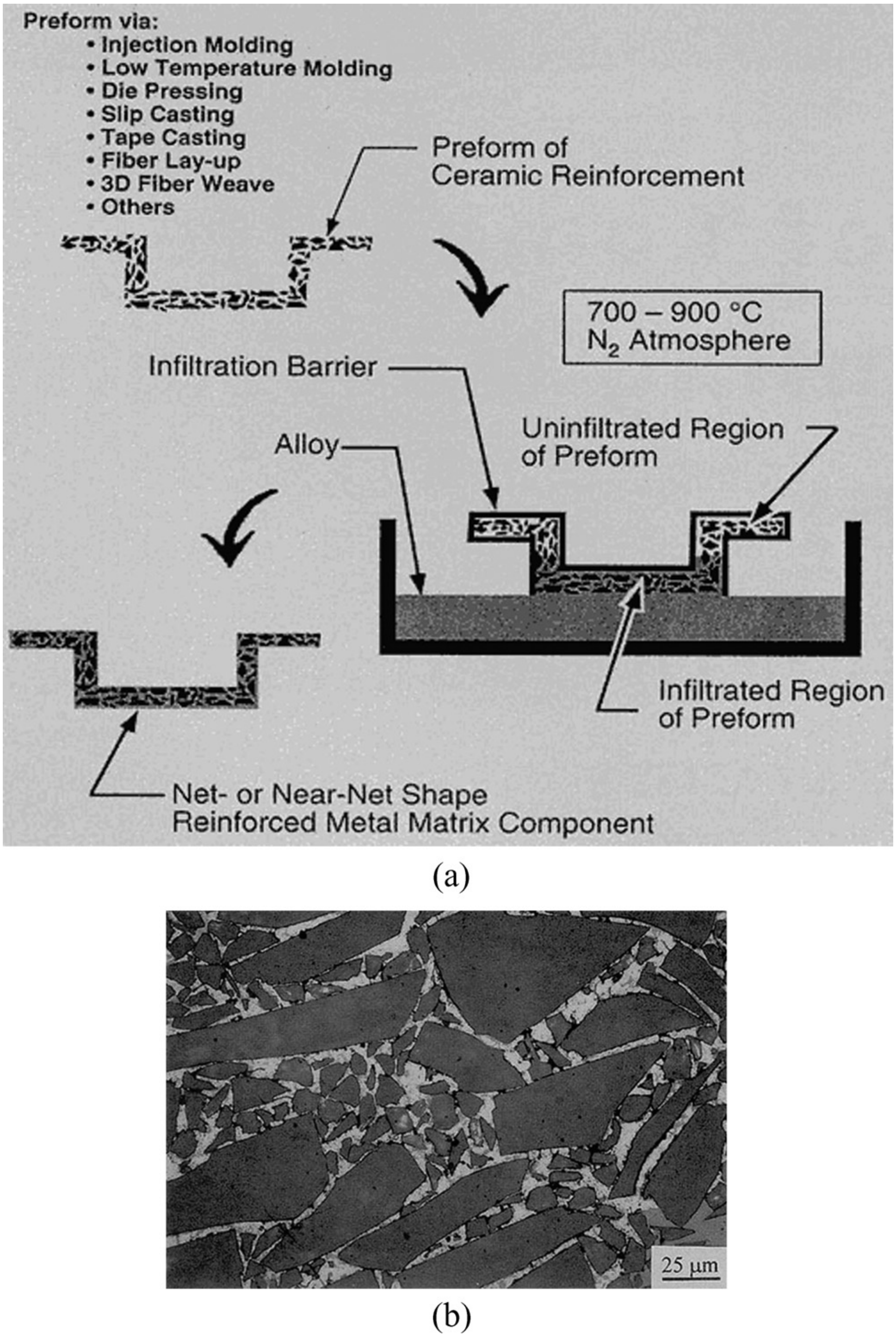

Preform via:
• Injection Molding
• Low Temperature Molding
• Die Pressing
• Slip Casting
• Tape Casting
• Fiber Lay-up
• 3D Fiber Weave
• Others
Preform of Ceramic Reinforcement
700 – 900 °C
N₂ Atmosphere
Infiltration Barrier
Uninfiltrated Region of Preform
Alloy
Infiltrated Region of Preform
Net- or Near-Net Shape Reinforced Metal Matrix Component
(a)
25 μm
(b)

Figure 2
PRIMEX process, showing (a) a schematic representation and (b) an optical micrograph of an Al–70 vol.% SiC MMC (Clyne 2000).

A special type of reactive processing is that in which a gaseous product is released, generating a metallic foam. These are often made from MMC stock material, since foam production requires raising the melt viscosity sufficiently to ensure that bubbles created by gas evolution do not collapse prematurely. Ceramic particles have this effect, which is illustrated by the foam structures shown in Figs. 3(a) and (b). However, the presence of (relatively coarse) ceramic particles within the cell walls can promote cell wall rupture (see Fig. 3(c)), leading to coarsening.

2.2 Solid-state Processing

(*a*) *Powder blending and consolidation*. Blending of metallic powder with ceramic fibers or particulates is a versatile technique for MMC production. This is usually followed by cold compaction, canning, evacuation, degassing, and a high-temperature consolidation stage such as hot isostatic pressing (HIPing) or extrusion. Achieving a homogeneous mixture can be difficult, particularly with fibers. A feature of much powder route material is the presence of fine oxide particles, usually present in aluminum MMCs in the form of plate-like particles a few tens of nanometers thick, constituting about 0.05–0.5 vol.%, depending on powder history and processing conditions. This fine oxide tends to act as a dispersion-strengthening agent and often has a strong influence on the matrix properties, particularly at high temperature.

MMCs produced by powder blending are commonly *extruded*. This can generate alignment of fibers parallel to the extrusion axis, but often at the expense of progressive *fiber fragmentation*. The degree of fiber fracture decreases with increasing temperature and decreasing local strain rate. Other microstructural features of extruded MMCs include the formation of ceramic-enriched *bands* parallel to the extrusion axis. The mechanism of band formation is still unclear, but it appears to involve the concentration of shear strain in regions where ceramic particles or fibers accumulate. However, extrusion of consolidated MMCs, such as castings, can *reduce* the level of clustering and inhomogeneities in the material. In general, the microstructural changes occurring in MMCs during forming processes are explicable in terms of local temperatures, stresses, and strain rates. Processes such as *rolling* and *forging* involve high deviatoric strains being imposed quickly, and hence can cause damage such as cavitation, particle fracture, and macroscopic cracking, particularly at low temperature. Very high temperatures, and the possibility of matrix liquation, however, can cause macroscopic defects such as *hot tearing* or *hot shortness*. In contrast to these forming processes, HIPing generates no (volume-averaged) deviatoric stresses and so is unlikely to give rise to either microstructural or macroscopic defects. It is an attractive method for removing residual porosity, which can include surface-connected porosity as long as some form of encapsulation is provided. It has been quite widely applied to MMCs. However, it can be very difficult to remove residual porosity in regions of very high ceramic content, such as within particle clusters, and the absence of any macroscopic shear stresses means that such clusters are not readily dispersed during HIPing. Furthermore, the oxide skins that normally coat powder particles of aluminum or magnesium are usually left intact by HIPing, leading to poor interparticle bonding.

(*b*) *Diffusion bonding of foils*. Monofilament-reinforced MMCs are mainly produced by the foil–fiber–foil (diffusion bonding) route or by the evaporation of relatively thick layers of matrix material onto the surface of the fiber (see Sect.2.3(a)). Work in this area is very much oriented towards titanium and titanium-based matrices. Fortunately, titanium diffusion bonds to itself very readily, mainly because it dissolves its own surface oxide layer at elevated temperatures in controlled atmospheres. Titanium reinforced with long fibers is commercially produced by the placement of arrays of fibers between thin metallic foils, often involving a *filament winding* operation, followed by *hot pressing*. One of the main problems lies in avoiding excessive chemical reaction at the fiber/metal interface. Additions of aluminum, molybdenum, or vanadium slow the kinetics of interfacial reaction, but also tend to make the rolling of thin foils more difficult. There is interest in *fiber coatings* designed to reduce these problems of interfacial attack (see Sect. 1.3). In general, the foil–fiber–foil route is cumbersome and obtaining high fiber volume fractions and homogeneous fiber distributions is difficult unless special techniques are used. Also, the process becomes difficult when the objective is to produce parts of complex shape. Diffusion bonding has also been used to fabricate laminates composed of alternate layers of metal and ceramic.

2.3 Vapor-state Processing

(*a*) *Physical vapor deposition* (*PVD*). All PVD processes are relatively slow, but the fastest is *evaporation*—involving *thermal vaporization* of the target species in a vacuum. An evaporation process used for fabrication of monofilament-reinforced titanium involves passing the fiber through a region having a high vapor pressure of the metal to be deposited, where condensation takes place so as to produce a thick surface coating. The vapor is produced by directing a high-power (~10 kW) electron beam onto the end of a solid bar feedstock. Alloy composition can be tailored, since differences in evaporation

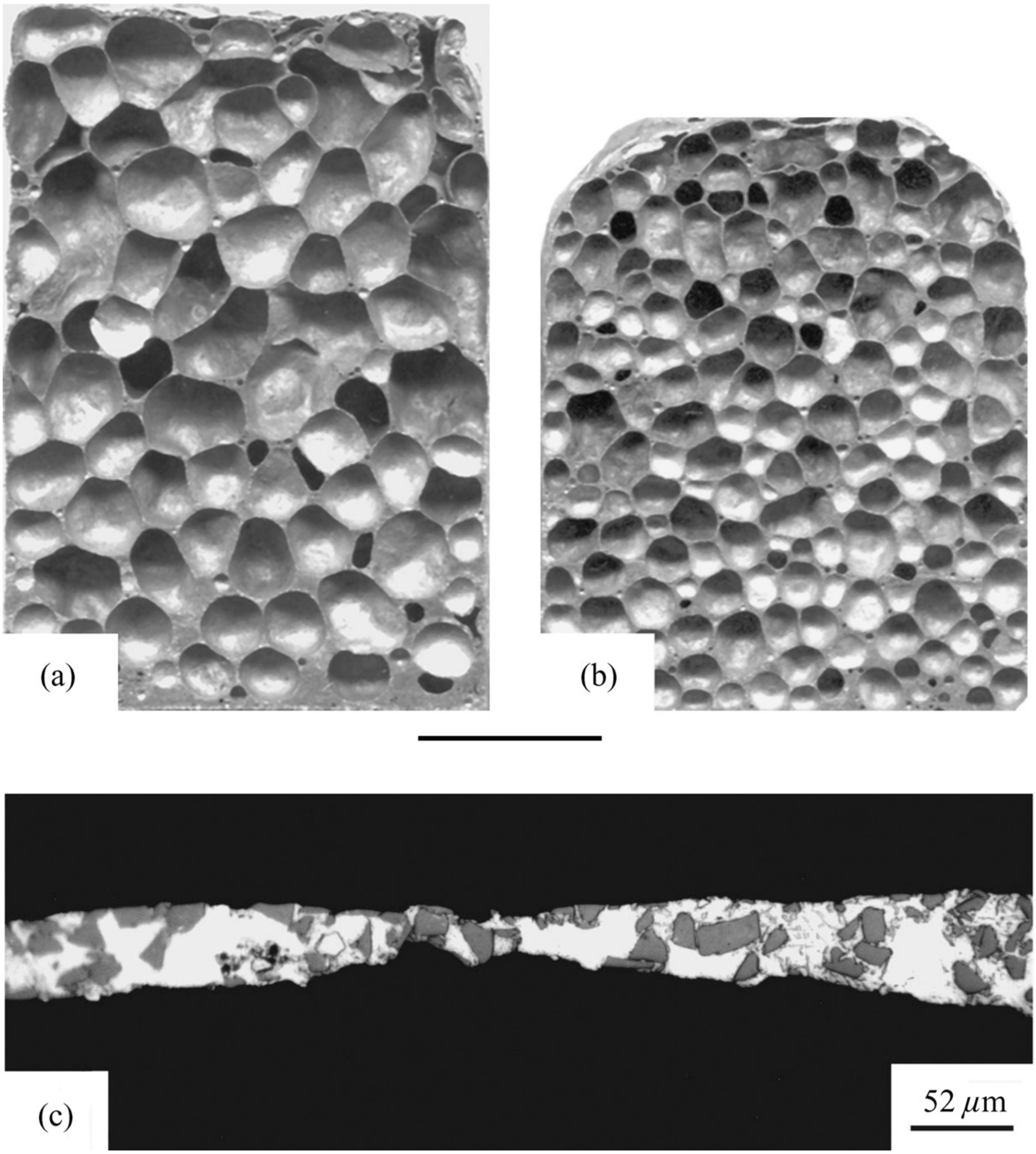


Figure 3
Structures of metallic foams produced by a melt route. Two cell structures are shown, corresponding to the same production conditions except that the SiC particulate content is (a) 10 wt.% and (b) 20 wt.%. (c) Microstructure of a cell wall for a 20 wt.% SiC_p case (Gergely *et al.* 2000).

rate between different solutes become compensated for by changes in composition of the molten pool formed on the end of the bar, until a steady state is reached in which the alloy content of the deposit is the same as that of the feedstock. A further point worth noting is that there is little or no mechanical disturbance of the interfacial region; this may be of significance when the fibers have a diffusion barrier layer, or a tailored surface chemistry. Typical deposition rates are $\sim 5\text{–}10\,\mu\mathrm{m\,min}^{-1}$. Composite fabrication is usually completed by assembling the coated fibers into a bundle and consolidating this in a hot pressing or HIPing operation. A very uniform distribution of fibers is produced in this way, with fiber contents of up to about 80%. A typical coated fiber and consolidated titanium composite are shown in Fig. 4. The fiber volume fraction can be accurately controlled via the thickness of the deposited coatings and the fiber distribution is always very homogeneous.

3. *Summary*

MMCs represent a wide range of materials, which includes cermets and metallic foams, as well as more conventional particle- and fiber-reinforced metals. Techniques employed for production of MMC material and components depend on the types of matrix and reinforcement concerned; these are classified according to whether the matrix is in the liquid, solid, or gaseous state when it is combined with the reinforcement. Each of these processing routes has advantages and disadvantages. In particular, some

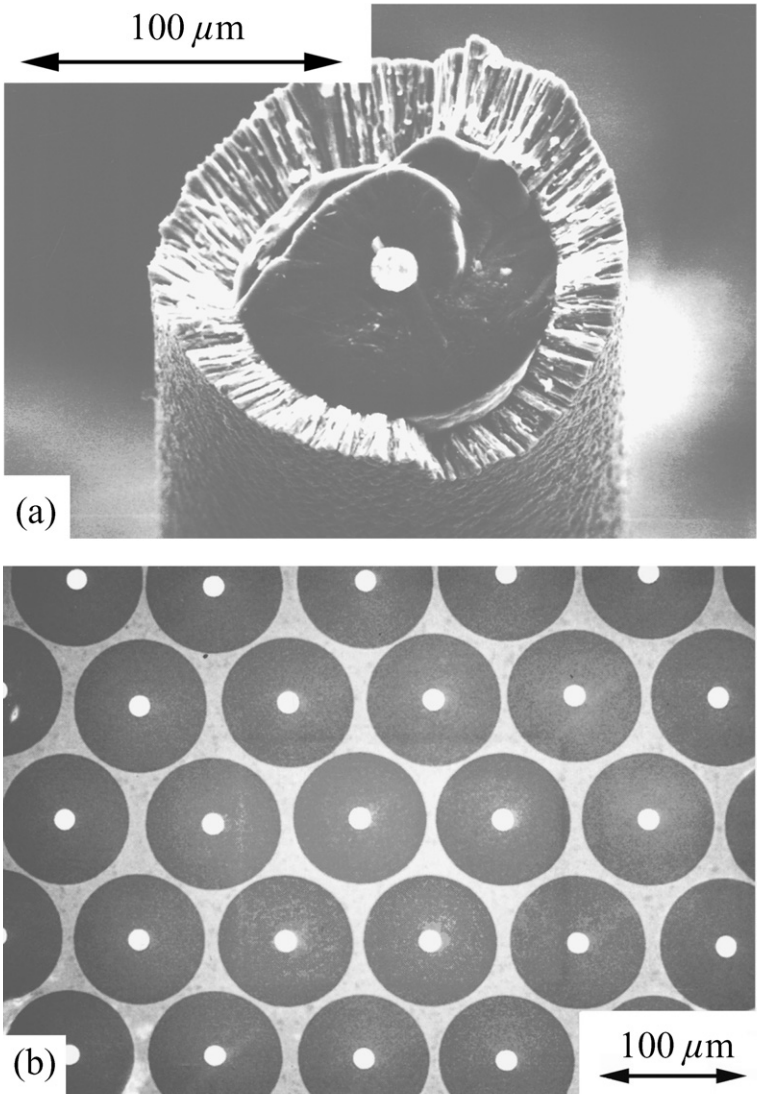


Figure 4
(a) SiC monofilament with a 35 μm vapor-deposited layer of Ti–5Al–5V and (b) Ti–5Al–5V/80 vol.% SiC composite produced by HIPing of a bundle of monofilaments with 8 μm thick coatings (Ward-Close and Partridge 1990).

are far more expensive than others. The lowest cost routes are generally those in which particle-reinforced aluminum is produced using liquid metal handling—particularly stir casting. Material produced in this way represents a substantial proportion of the MMCs in commercial use. Other materials and processes have, however, also gained in importance. For example, melt infiltration techniques are used to produce components for automotive engine and electronic substrate applications and powder processes are employed for aerospace applications. It is likely that, as MMC applications continue to expand, the range of materials and processes employed will remain relatively wide.

Bibliography

Aghajanian M K, Rocazella M A, Burke J T, Keck S D 1991 The fabrication of metal matrix composites by a pressureless infiltration technique. *J. Mater. Sci.* **26**, 447–54

Clyne T W 1996 Metallic composite materials. In: Cahn R W, Haasen P (eds.) *Physical Metallurgy*. Elsevier, pp. 2568–625

Clyne T W 2000 An introductory overview of MMC systems, types and developments. In: Clyne T W (ed.) *Comprehensive Composite Materials, Vol. 3. Metal Matrix Composites*. Elsevier, Amsterdam, pp. 1–26

Clyne T W, Withers P J 1993 *An Introduction to Metal Matrix Composites*. Cambridge University Press, Cambridge, UK

Everett R K 1991 Deposition technologies for MMC fabrication. In: Everett R K, Arsenault R J (eds.) *Metal Matrix Composites: Processing and Interfaces*. Academic Press, New York, pp. 103–19

Georghe I, Rack H J 2000 Powder processing of metal matrix composites. In: Clyne T W (ed.) *Comprehensive Composite Materials, Vol. 3. Metal Matrix Composites*. Elsevier, Amsterdam, pp. 679–700

Gergely V, Degischer H P, Clyne T W 2000 Recycling of MMCs and production of metallic foams. In: Clyne T W (ed.) *Comprehensive Composite Materials, Vol. 3. Metal Matrix Composites*. Elsevier, Amsterdam, pp. 797–820

Li B, Lavernia E J 2000 Spray forming of MMCs. In: Clyne T W (ed.) *Comprehensive Composite Materials, Vol. 3. Metal Matrix Composites*. Elsevier, Amsterdam, pp. 617–53

Lloyd D J 1994 Particle reinforced aluminium and magnesium matrix composites. *Int. Mater. Rev.* **39**, 1–23

Lloyd D J, Jin I 2000 Melt processed aluminium matrix particle reinforced composites. In: Clyne T W (ed.) *Comprehensive Composite Materials, Vol. 3. Metal Matrix Composites*. Elsevier, Amsterdam, pp. 555–77

Mortensen A 2000 Melt infiltration of metal matrix composites. In: Clyne T W (ed.) *Comprehensive Composite Materials, Vol. 3. Metal Matrix Composites*. Elsevier, Amsterdam, pp. 521–54

Mortensen A, Jin I 1992 Solidification processing of metal matrix composites. *Int. Mater. Rev.* **37**, 101–28

Mortensen A, Masur L J, Cornie J A, Flemings M C 1989 Infiltration of a fibrous preform by a pure metal. *Metall. Trans.* **20A**, 2535–63

Ward-Close C M, Partridge P G 1990 A fiber coating process for advanced metal matrix composites. *J. Mater. Sci.* **25**, 4315–23

Ward-Close C M, Robertson J G, Godfrey S P 2000 Fabrication of monofilament-reinforced titanium. In: Clyne T W (ed.) *Comprehensive Composite Materials, Vol. 3. Metal Matrix Composites*. Elsevier, Amsterdam, pp. 655–78

T. W. Clyne
University of Cambridge, UK

Metallic Filaments

Thin wires or "filaments" of metal, by definition of diameter below about half a millimeter, have been used for many years in a variety of applications. Besides many well-established applications, such as tungsten filaments in the production of electric lightbulbs or thin wires used in electronics, thin metallic wires or filaments also represent a particular class of continuous high-strength fibers that can be used as reinforcements in composite materials. This article gives a short overview on the production methods, the particular properties, and some typical applications of thin wires and metallic filaments.

1. Drawn Wires and Filaments

Metallic filaments of submillimetric dimensions are produced by various techniques. The most frequently used method is the classical wire drawing technique (Mielnik 1991), in which wires are successively pulled through thinner and thinner dies (Fig. 1). Clearly this technique works best with ductile metals such as copper. Less ductile metals have to be drawn at elevated temperature. Metals with a high work

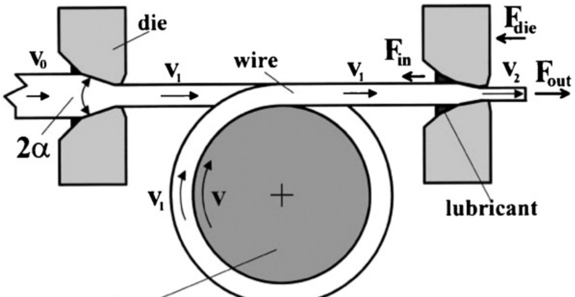


Figure 1
Drawing in a multiple-die machine. The drawing speeds *v* increase after each diameter reduction and must be closely controlled or slippage must be allowed to guarantee continuity. Drawing speeds of up to several meters per second are common.

hardening rate need more frequent annealing steps in this multistage process.

Wires are typically drawn with a drawing force that corresponds to 60% of their ultimate tensile strength. This force equals the sum of the plastic deformation force and the force needed to overcome friction in the die. For a given diameter reduction, the friction is larger than the deformation force in dies with low opening angles, whereas in dies of larger opening angles the deformation force becomes dominant. The minimum drawing force, which also depends on the metal to be drawn, its state of work hardening, and the lubricant used, occurs when the total opening angle 2α of the die is in the range 10–20°. With an appropriately designed die such forces allow a reduction of the cross-section of 20–40% per die for ductile metals and 10–20% for less ductile ones. Bonding wires for applications in electronics, made of pure gold, copper, and aluminum, with diameters from 20 to 50 μm, are typically produced from a wire of 1 mm initial diameter. Consequently they need about 30–40 steps before they reach their final dimensions. For less ductile metals this number easily doubles. Further reductions of the diameter down to a few microns is still possible with this technique but practical and fundamental limits of this process rapidly become important. Due to the relative increase of the wire surface area, friction becomes more and more important with respect to the drawing force and, consequently, imposes smaller diameter reductions per step. Finally, towards the smallest diameters, safely manipulating wires of the thickness of spider silk and having a breaking load of less than 1 g becomes the dominant problem.

Since diamond has the highest resistance to abrasion, dies of natural monocrystalline or synthetic polycrystalline diamonds are the preferred tools for the production of fine and ultrafine wires. In order to limit the consumption of diamonds, used dies are reshaped and reinserted in a preceding step. Dies made of tungsten carbide and other hard metals are mainly used for wires and rods of larger diameters.

Practically all steels can be drawn to thin wires. For applications where high tensile strength is demanded, plain carbon steels with 0.5–0.9% carbon are well suited. The strength of steel wires is mainly controlled by the steel carbon content and its microstructure, which in turn can be modified by heat treatments and the work hardening that occurs during the drawing process. Typically hot rolled rods are cooled by forced air at a much lower cooling rate (5–30 $^\circ\mathrm{C\,s}^{-1}$) than by quenching in water. This gives grains of intermediate sizes which give sufficient ductility to allow cold drawing down to the desired diameters. The extreme cold work gives rise to a considerable strain hardening, which raises the tensile strength well beyond the level that could be achieved by quenching. The strongest carbon steel wires (0.9% C), also known as piano wires (used as piano cords), achieve a tensile strength of over 3 GPa by this treatment. Major applications of such steel wires are the reinforcement of tires, and fabrication of stranded cables for flexible force transmission, cable cars, and suspension bridges. In order to get improved corrosion resistance, stainless steel wires are used. They often have a low carbon content, which results in a lower tensile strength but higher ductility. Since work hardening is as effective in stainless steels as in carbon steels, appreciable tensile strengths can also be achieved. In some high alloy steels the ultimate tensile strength can be further increased after drawing by precipitation hardening. Stainless steel wires with diameters as low as 10 μm are commercially available. A major application of thin stainless steel wires is in the fabrication of sieves.

Bonding wires are typically made from high-purity gold. More recently pure copper and aluminum–1% silicon have become also competitive substitutes. Very small amounts of other metals may be added to modify their mechanical performance. These wires are used to electrically connect integrated circuits on silicon chips to the metallic pins and are selected according to the bond strength that results following thermal compression or ultrasonic compression bonding on the metallized silicon surface and the pins.

Bonding wires, as well as wires made from other ductile metals, which are submitted to substantial reductions of their cross-section without intermediate recrystallization annealings, often develop extremely strong textures. Drawn gold and copper wires that have undergone a diameter reduction from 2 mm to 50 μm without intermediate and final heat treatment, for instance, develop an almost perfect (100)-texture (100) axis parallel to the drawing direction. In this state their Young's modulus along the wire axis approaches the modulus measured along the (100)-direction in a single crystal. For copper this gives an observed modulus of the textured wire of 65 GPa whereas the modulus $E(100)$ for a single crystal is 60 GPa and the isotropic modulus in a polycrystal amounts to 120 GPa.

The use of fine tungsten filaments in incandescent lamps is an application that dates back to the nineteenth century. Tungsten has important advantages over filaments made of carbon or other metals. It not only has the highest melting point, but also the lowest vapor pressure of all metals. Moreover its mechanical strength even at high temperatures remains satisfactory and its emissivity in the visible part of the spectrum is particularly favorable. Tungsten wires are produced from sintered ingots of tungsten powder containing grain growth inhibitors. These ingots are first swaged and then drawn to the desired diameter. The diameter and length of the filaments operating at a given temperature depends on the electrical tension and the desired power of the lamp. For a standard lamp (220 V and 100 W) filaments of about 50 μm diameter and 1 m length are needed. In order to accommodate this length in a bulb of usual size the filaments are doubly coiled, first on small mandrel (0.1 mm) and then on bigger one (0.5 mm). Over the years many metallurgical studies have contributed to improve the lifetime of tungsten lamps. The major problems are grain growth in the hot filaments, which may lead to the so-called bamboo structure and evaporation. Since tungsten filaments are operated near or above 2000 °C and the recrystallization temperature is only about 1200 °C, grain growth, which makes the filaments extremely brittle, must be prevented. By adding 1–2% of a very fine powder of ThO_2 or by doping with small amounts (few ppm) of alumina or potassium silicate, followed by intense deformation, it is possible to produce a tungsten in which the small and elongated grains typical of drawn wires remain stable. The problem of evaporation has no metallurgical solution. It is tackled by filling the bulb with argon (about 0.8 atm) or, as in halogen lamps, by employing a regenerative chemical cycle inside the glass that redeposits the evaporated tungsten on the filament.

2. The Wollastone Technique

This method is particularly useful to draw wires of a few microns diameter or to draw brittle metals. The metal to be drawn is tightly packed in a tube made of a ductile metal. Both metals are then drawn by the conventional drawing technique until the desired reduction is achieved. To get the fine wire the surrounding metal is removed chemically. In several important cases, however, this separation is not necessary since the resulting composite forms the desired end-product. In particular, this is the standard method to produce superconducting wires as well as thermocouples and heating wires with metallic sleeves. For example, most of the technically interesting superconducting alloys, e.g., NbTi and Nb_3Sn, are extremely brittle and cannot be drawn directly to form very thin wires, whereas when packed in a copper or brass rod drawing becomes possible. For reasons related to the stability of the superconducting state such conductors are usually formed by hundreds or thousands of individual filaments embedded in the matrix, rather than by one single wire.

Coaxial thermocouples and heating wires are produced by the same technique. In order to isolate electrically the external sleeve from the inner wires, very fine powders of oxides (e.g., MgO) are packed in the initial rod before drawing. Compared to an electrical insulation made from sintered ceramic capillaries and tubes the coaxial wires and couples have the advantage of remaining flexible.

3. The Taylor Method

This drawing technique takes advantage of the ease of production of glass fibers by drawing from the molten end of a solid glass bar. The Taylor technique (Nixdorf 1970) uses the same method on glass tubes containing a core of the molten metal to be formed. By selecting a glass with a softening temperature in the vicinity of the melting temperature of the metal to be formed in a filament, the enclosed liquid metal will follow with practically no friction to the imposed cross-section reduction of the continuously deforming glass tube. Clearly, there must be good chemical compatibility between the glass and the metal. Finally, the capillary sheet surrounding the metallic filament has to be removed; this can be achieved using HF for noble metals and stainless steel, or alternatively molten alkaline salts, which also dissolve glasses of any composition and show little reaction with many metals. The Taylor method can be used to produce wires of many common metals and alloys down to 1 μm.

4. Metallic Filaments Directly Produced from the Melt

It is quite easy to produce a cylindrical jet of liquid metals. Unfortunately, the combination of high surface tension and low viscosity in liquid metals leads to a rapid disintegration into small droplets (Rayleigh instability). To produce thin wires directly from the melt, the liquid metal therefore has to be cooled and solidified within a very short distance (<1 cm) from the orifice. This is the principle used in the melt spinning method (cooling on a solid substrate) and the in-rotating-water spinning technique (liquid quenching medium).

The meltspinning technique (Cahn 1983) is mainly used to produce 20–60 μm thick ribbons of glassy metals. A typical melt spinning apparatus consists of a rotating wheel, an induction heater and a crucible with a round or rectangular nozzle at its bottom (Fig. 2). The width of the nozzle essentially defines the width of the ribbon to be produced. The distance between the nozzle and the wheel surface is often adjusted in such a manner that a sessile drop forms between the moving wheel and the orifice. After the metal is molten, the wheel is set into rotation to a surface velocity of about $30\,\mathrm{m\,s^{-1}}$ ($110\,\mathrm{km\,h^{-1}}$) and the liquid metal is ejected through the orifice by slightly raising the pressure inside the crucible. The liquid puddle that forms on the wheel surface immediately starts to cool and a solidification front moves into the droplet. Since the solidified metal, which forms the ribbon, moves with the wheel this front remains spatially stationary. The ribbon then remains in contact with the wheel until it detaches after about 10 cm. This method has been used to produce ribbons

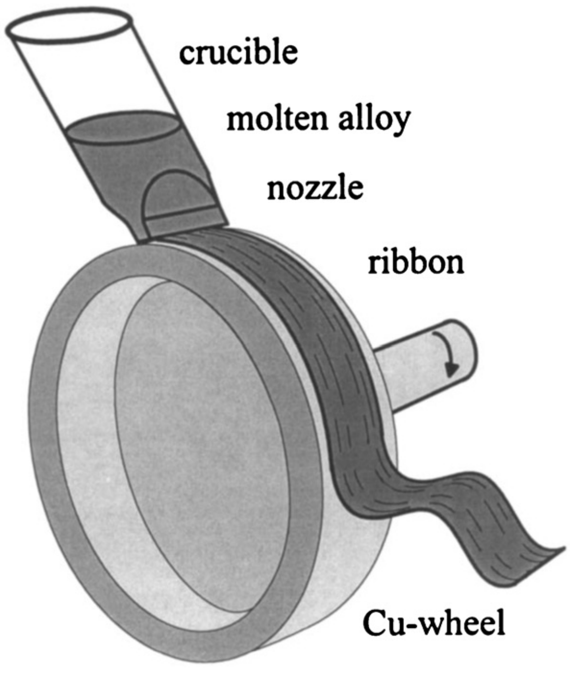


Figure 2
Melt spinning apparatus.

and fibers of amorphous alloys of many different compositions (e.g., $Fe_{80}B_{20}$, $Pd_{80}Si_{20}$, $Fe_{40}Ni_{40}B_{20}$).

Several amorphous alloys are nowadays industrially produced by this method for magnetic and mechanical applications. By machining equally spaced transversal grooves of about 1 mm depth on the wheel, fibers of a fixed length can also be produced. Some fibers made from iron-based alloys may have a tensile strength of over 3 GPa; these have been used to reinforce concrete. Ribbons and fibers produced by this technique unfortunately often have serrated edges and mold-like impressions, which stem from air turbulence and air inclusions between the wheel surface and the liquid metal. Such surface defects reduce the intrinsically high tensile strength of the amorphous metals, which is why melt spinning is often practiced in a vacuum to reduce or eliminate these defects.

Improved surface quality and circular wire cross-sections can be achieved by quenching the jet of molten metal into a liquid medium. The "in rotating water technique" uses water that is kept within a groove on the inner surface of a hollow rotating wheel Hagiwara *et al.* (1982). Contrary to melt spinning, the major problem with this technique is that it is very intolerant to even small variations of many casting parameters such as angle of jet incidence, or absolute and relative velocities of the rotating water and jet.

5. *Quantum Wires*

The progressive miniaturization of electronic devices also calls for thinner and thinner conductors. Submicron filaments can be produced by optical, x-ray, and electronic lithography. Such filaments, however, cannot be handled without the substrate on which they have been produced. The production of even smaller filaments, the so-called quantum wires, is presently a field of intense research. Their name stems from the fact that due to their small diameter, quantum mechanical effects in the electronic transport properties become important. In order to reach this limit their lateral dimension should be no greater than a few atomic diameters (<50). Presently such filaments have been produced by field evaporation of silver from the tip of a scanning tunneling microscope.

6. *Conclusion*

Drawing of wires and rods is an ancient technology. Its origin dates back to 3000 BC in the Middle East (Blake-Coleman 1992). Continuous drawing machines came along with the Industrial Revolution and were subsequently developed to near perfection. Nowadays the drawing of wires is clearly the most important manufacturing process for metallic filaments. The Wollastone technique allows the production of composite wires that could hardly be made by other techniques. The Taylor method has so far gained little or no industrial importance. The use of the melt spinning technique for the production of amorphous metallic filaments is more recent but has reached industrial scale. Due to their excellent mechanical properties, metallic filaments have a good potential for use as reinforcement fibers in composites for applications in which the relatively high density of metals carries little, or no, penalty.

Bibliography

Blake-Coleman B C 1992 *Copper Wire and Electrical Conductors—The Shaping of a Technology*. Harwood Academic, London

Cahn R W 1983 Alloys rapidly quenched from the melt. In: Cahn R W, Haasen P (eds.) *Physical Metallurgy*, 3rd edn. Elsevier, Amsterdam

Hagiwara M, Inoue A, Msumoto T 1982 Mechanical properties of Fe–B–Si amorphous wires produced by in-rotating-water spinning method. *Metall. Trans. Ser. A* **13A**, 373–82

Mielnik E M 1991 *Metallworking Science and Engineering*. McGraw-Hill, New York

Nixdorf J 1970 Comparison between whiskers and filaments. *Proc. Roy. Soc. London* **319**, 17–32

H. U. Künzi
EPFL, Lausanne, Switzerland

Mineral-filled Polymers

Particulate-filled polymers are widely used in many fields of application. With the exception of poly(vinyl chloride) (PVC) it is mostly technological reasons that justify the introduction of fillers, since a separate compounding step increases considerably the price of the composite. The combination of a filler with a polymer results in a new material with changed properties. Some of these changes are advantageous (improved stiffness, higher heat deflection temperature, better heat conductivity) while others are less favorable (decreased deformability and impact toughness, wear of the processing equipment). In this article a brief account is given of the most important fillers used in plastics, the method of incorporation, and the factors influencing the properties of the composites; finally, some areas are mentioned where mineral filled polymers are used in considerable quantities.

1. *Fillers*

The most important particulate fillers are described below; more detailed information can be found in Katz and Milewski (1978), Wypych (1999), and Rothon (1995).

1.1 Calcium Carbonate

The application of $CaCO_3$ in polymers is 10-fold greater than the use of any other filler. This mineral exists in three crystalline modifications—calcite, aragonite, and vaterite—but only calcite has practical importance. It can be found in large quantities all over the world, but minerals mined at different locations differ considerably in purity, size of the crystals, and origin, which all influence their use as fillers. Calcium carbonate is also prepared via synthetic routes; however, the price of precipitated $CaCO_3$ is considerably higher than that of the ground mineral, and thus it can be used only in special applications.

1.2 Talc

The chemical composition of pure talc is $Mg_3(Si_4O_{10})(OH)_2$; it is a secondary metamorphic mineral. The filler consists of platelets in which the layers of the particles are bonded only by secondary van der Waals forces. Cleavage takes place very easily under the conditions of usual thermoplastic processing technologies. Talc is the softest mineral with a Mohs hardness of 1. Its surface is rather hydrophobic; active hydroxy groups can be found only on broken edges and surfaces, thus the surface treatment of the filler is difficult. It reinforces polymers due to its platelet-like geometry and has a strong nucleation effect in some crystalline polymers, further increasing its stiffening effect.

1.3 Glass Fiber

Short glass fibers represent a transition between anisotropic mineral fillers and long fiber reinforcements. They are introduced mostly into engineering thermoplastics, but considerable quantities are used also in polypropylene (PP) and in sheet molding compounds. Short carbon-fiber reinforced composites have also appeared on the market. The proper surface treatment of the fiber is crucial in all fiber-reinforced composites.

1.4 Other Fillers

Kaolin is used in various forms in polymers (china clay, calcined clay). The anisotropic particle geometry of wollastonite and mica reinforces the polymer. Barite-filled polymers, mostly PP, have excellent vibration-damping properties due to the high density of this filler. Wood flour and natural fibers (flax, hemp, jute, sisal, etc.) are applied in increasing quantities, mostly in composites for the automotive industry. The water absorption of these fillers creates problems in processing and application. Increasing interest is being shown in layered silicates, mostly in montmorillonite. The nanocomposites prepared from these minerals possess unique properties at reduced filler content. The main characteristics of some mineral fillers are listed in Table 1 (see also *Clay-based Polymer Nanocomposites**).

2. *Compounding*

The polymer and the filler must be homogenized prior to the manufacture of a product, since insufficient homogenization leads to aggregation. Aggregates cause surface blemishes and affect adversely all the properties, especially impact strength, because they initiate cracks under external loading (see *Polymer Melt Mixing: Agglemarate Dispersion**).

Compounding is mostly done in twin screw extruders in a continuous operation, with high productivity and flexibility. The machine configuration and processing conditions can be adjusted according to the properties of the processed material. High shear must be used for the homogenization of particulate-filled polymers to avoid aggregation, but low shear and intensive mixing is preferred in the case of shear-sensitive polymers (e.g., PVC) and fillers (e.g., starch), as well as glass fibers. The attrition of fibers leads to the loss of their reinforcing effect and generally to the deterioration of composite properties.

3. *Factors Influencing the Properties of Filled Polymers*

The properties of mineral-filled polymers are determined by the characteristics of the components,

Table 1
Composition and properties of particulate fillers.

Filler	Composition	Density ($g\,cm^{-3}$)	Mohs hardness	Shape
Calcium carbonate	$CaCO_3$	2.7	3	sphere
Talc	$Mg_3(Si_4O_{10})(OH)_2$	2.8	1	plate
Kaolin	$Al_2O_3 \cdot 2SiO_2 \cdot 2H_2O$	2.6	2.5–3.0	plate
Wollastonite	$CaSiO_3$	2.9	4.5	needle
Mica	$KM(AlSi_3O_{10})(OH)_2$	2.8	2.0–2.5	plate
Barite	$BaSO_4$	4.5	3.5	plate

composition, interfacial interactions, and structure. All four factors can vary in a wide range, thus allowing the preparation of tailor-made composites.

3.1 Component Properties

Important properties of the matrix polymer are its chemical composition and melt viscosity. The first determines the strength of interaction with a nontreated filler, while the second influences the shear stresses developing during homogenization. The extent of aggregation, as well as the smallest particle size that can be used in a polymer, are determined by the relative magnitude of adhesion and shear forces acting in the melt during homogenization (Pukánszky and Fekete 1999). The mechanical properties of the matrix polymer strongly influence the micromechanical deformation processes taking place under the effect of external load. Large stresses develop in a stiff matrix, and the dominating deformation mechanism is usually debonding. In such cases yield stress and strength decrease with increasing filler content (Fig. 1). In a softer polymer (e.g., low-density polyethylene (LDPE)) the introduction of the same filler results in true reinforcement, with yield stress and strength increasing with increasing amount of filler.

Numerous filler characteristics influence the properties of the composites. Chemical composition and purity affect color and stability. However, mechanical properties and performance are determined mainly by particle size, particle size distribution, specific surface area, and shape. An optimum particle size exists for each polymer and processing technique. Large particles debond easily from the matrix under loading, while small particles tend to aggregate; both processes weaken the composite properties significantly. The specific surface area determines the contact surface between the filler and the polymer and thus the amount of interphase formed. The surface free energy (surface tension), hardness, and thermal properties of the filler also influence the production or application of mineral-filled polymer composite parts.

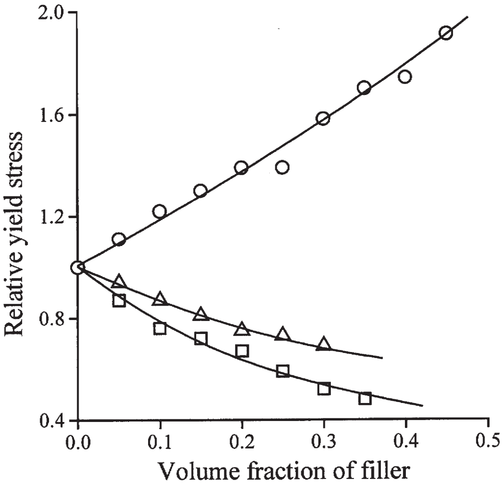


Figure 1
Effect of matrix properties on the reinforcing effect of a particulate filler ($CaCO_3$, 3.6 μm): ○ LDPE; △ PP; □ PVC.

3.2 Composition

The stiffness of the composite increases, while its price often decreases, with increasing filler content; therefore, significant effort is concentrated on the introduction of the highest possible filler loadings into the polymer. Since other properties deteriorate at the same time, a composition must be found where the combination of all properties is optimized.

It is generally held that tensile yield stress and tensile strength decrease with increasing filler content when fillers of nearly spherical shape are used. The composition dependence of these properties is also strongly influenced by particle characteristics. As Fig. 1 shows, true reinforcement can be achieved with fillers in a soft polymer, and if the particle size is sufficiently reduced the same effect is observed in PP (see Fig. 2) and also in PVC (Pukánszky 1995). The effect of anisotropic fillers depends on the orientation of the particles. Deformability, yield strain,

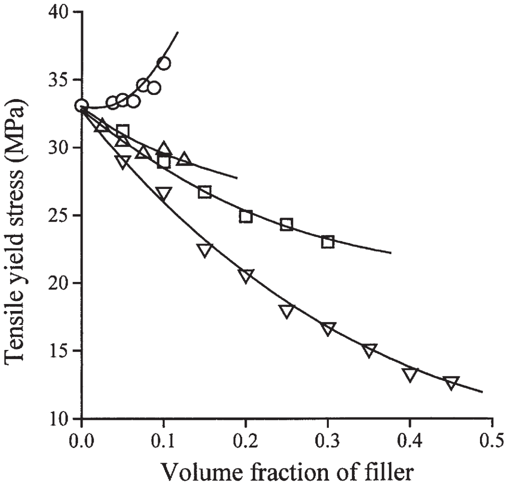


Figure 2
Effect of particle size on the tensile yield stress of PP composites. Particle diameter: ○ 0.01 µm; △ 0.08 µm; □ 3.3 µm; ▽ 58 µm.

and elongation-at-break always decrease with increasing filler content. Impact toughness often exhibits a maximum at a certain filler content, which depends on the characteristics of the filler and on interfacial interactions (Pukánszky and Maurer 1995). The maximum is a consequence of several competitive effects; the debonding of the filler during deformation leads to energy consumption, while increasing filler loading leads to increased stiffness and decreased impact toughness. Most other properties (heat conductivity, heat deflection temperature, hardness) are more or less proportional to the filler content of the composite.

3.3 Interfacial Interaction

This factor strongly influences the properties of polymer composites. Although it is mostly weak van der Waals forces that act between the polymer and the filler in thermoplastics, they are sufficiently strong to bond the two components. The adsorption of the polymer on to the filler surface leads to the formation of an interphase region in which molecules have decreased mobility. The effect of interaction depends on the size of the contact surfaces (specific surface area of the filler) and the strength of the interaction. Smaller particles have a larger surface area leading to increased yield stress (Fig. 2).

To modify interfacial interactions four types of methods are frequently used. *Nonreactive treatment* (surfactants) leads to a decrease in particle–particle and matrix–filler interaction. The latter results in decreased strength and increased deformability, and occasionally to increased impact resistance, while the former leads to improved processability, homogeneity, and appearance. *Reactive treatment* (coupling), which assumes covalent coupling of both components, is difficult in most thermoplastics due to the lack of reactive groups on the polymer. Reactive coupling is easier in thermoset matrices. The introduction of *functionalized polymers* into the composite was shown to improve adhesion and increase yield stress as well as tensile strength in the case of various fillers, from $CaCO_3$ through wood flour and magnesium hydroxide to glass fibers (Fig. 3) (Felix and Gatenholm 1993). This technology is used mostly in polyolefins, and especially in PP, but occasionally also in other polymers. The fourth method of modification involves multicomponent polymer systems containing a filler and two polymers. In these systems a special morphology may form in which one of the polymer components encapsulates the filler. Such a structure is accompanied by significantly modified interaction and properties (Pukánszky 1995).

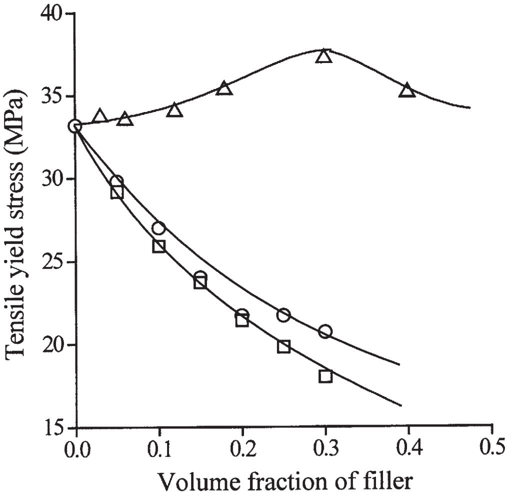


Figure 3
Modification of interfacial interaction in mineral filled composites. Effect on the tensile yield stress of PP/$CaCO_3$ (3.6 µm) composites: ○ nontreated; △ increased interaction (MAPP); □ decreased interaction (stearic acid).

3.4 Structure

Particulate-filled polymers are believed to have a simple structure consisting of uniformly dispersed particles in a homogeneous matrix. Two important structural phenomena must always be kept in mind when filled polymers are prepared and applied. Small particles tend to form aggregates which weaken the properties. Above a certain specific surface area the aggregation tendency of spherical nontreated fillers becomes significant. This critical surface area depends on several factors including the viscosity of the polymer, characteristics of the filler, processing technology, etc. The orientation of the anisotropic particles determines their reinforcing effect.

4. *Application*

Large amounts of $CaCO_3$ are used in PVC, both in plasticized (cable, flooring) and rigid (pipes, fittings, profiles) products. Although at a much lower level, $CaCO_3$ and also talc are often introduced into PP. Such composites are used in the automotive industry as bumpers, air-filter covers, or heater boxes, and as household articles. Artificial paper and breathable films are prepared from HDPE containing considerable amounts of $CaCO_3$. Sheet and bulk molding compounds represent the most important application in thermosets, and products prepared from them are used in construction, automotives, etc. The new class of mineral-filled polymers, nanocomposites, may be used mainly in structural applications. These composites possess improved mechanical properties, decreased gas permeation, and flame-retardation and the enhanced properties are accompanied by considerable weight savings.

5. *Summary*

The use of mineral fillers increases continuously. The incorporation of the filler into the polymer results in new materials. The properties of these composites are determined by the characteristics of the components, composition, interfacial interaction, and structure. The last factor is often neglected during the development of mineral-filled polymers, although both aggregation and orientation of anisotropic filler particles affect strongly the composite properties. Mineral-filled polymers are used in a wide range of applications, mostly in fields where increased stiffness and heat resistance are required. Although $CaCO_3$, talc, and glass fibers are used in the largest quantities, new fillers (natural fibers) and novel composites (nanocomposites) are also gaining increasing importance in this field.

See also: Polymer Matrix Composites with Roughly Equiaxed Reinforcements: Microstructure and Mechanical Behavior*

Bibliography

Felix J M, Gatenholm P 1993 Formation of entanglements at brushlike interfaces in cellulose–polymer composites. *J. Appl. Polym. Sci.* **50**, 699–703

Katz H S, Milewski J V 1978 *Handbook of Fillers and Reinforcements for Plastics*. Van Nostrand, New York

Manas-Zloczower I, Tadmor Z 1994 *Mixing and Compounding of Polymers: Theory and Practice*. Hanser, Munich, Germany

Nielsen L E 1974 *Mechanical Properties of Polymers and Composites*. Dekker, New York

Pukánszky B 1995 Particulate filled polypropylene: structure and properties. In: Karger-Kocsis J (eds.) *Polypropylene: Structure, Blends and Composites*. Chapman and Hall, London, Vol. 3, p. 1

Pukánszky B, Fekete E 1999 Adhesion and surface modification. *Adv. Polym. Sci.* **139**, 109–53

Pukánszky B, Maurer F H J 1995 Composition dependence of the fracture toughness of heterogeneous polymer systems. *Polymer* **36**, 1617–25

Rothon R 1995 *Particulate-Filled Polymer Composites*. Longman, Harlow, UK

Wypych G 1999 *Handbook of fillers*. Chemical and Technical Publishers, Toronto, ON

B. Pukánszky
Technical University of Budapest, Hungary

This Page Intentionally Left Blank

N

Nanoscale Ceramic Composites

A new class of structural ceramic materials with superior mechanical properties is currently being developed. These materials are known as "ceramic nanocomposites" and are produced by combining a structural ceramic matrix, such as Al_2O_3, ZrO_2, or Si_3N_4 with a nano-sized second phase. The second phases typically used are "nanoparticles" which are 300 nm or less in size and added in quantities between 1 vol.% and 30 vol.% (Todd 1995). Ceramic nanocomposites have been reported to possess both room- and high-temperature properties which are greatly superior to those of the monolithic matrix in a variety of material systems, indicating that a "structural synergism" occurs between the matrix and second phase in these systems (Niihara and Nakahira 1988, Ding *et al.* 1993, Pezzotti and Sakai 1994). In this article, the development of the most widely studied ceramic nanocomposite, Al_2O_3–SiC, will be discussed. In particular, the processing steps required to fabricate Al_2O_3–SiC nanocomposites and a review of their microstructures will be given. Most importantly, the improvements in mechanical properties over monolithic Al_2O_3 will be presented, and the possible mechanisms responsible for these improvements will be discussed.

1. Processing and Microstructure of Al_2O_3–SiC Nanocomposites

Al_2O_3–SiC nanocomposites are produced using standard ceramic processing techniques. Matrix and second-phase powders are typically comilled in an organic medium to achieve a uniform dispersion of second phase throughout the matrix. After milling, the nanocomposite powder is dried and then densified using pressureless sintering or hot pressing. Relative to that of monolithic Al_2O_3, densification in the nanocomposite is strongly inhibited by the presence of particles on grain boundaries, which reduce the rate of grain-boundary diffusion. As a result, the sintering conditions necessary to achieve full densification are strongly dependent on the amount of second phase, with the sintering temperature increasing significantly with the second-phase volume fraction. While monolithic Al_2O_3 can be pressureless sintered to near-theoretical density at 1400 °C, a temperature of 1775 °C is necessary for full densification in a nanocomposite of Al_2O_3 containing 5 vol.% SiC (Stearns *et al.* 1992). In fact, no study has shown that full density can be achieved using pressureless sintering in a nanocomposite containing more than 5 vol.% SiC. The difficulties of densification in the nanocomposite can be alleviated by hot pressing. During hot pressing, an applied pressure increases the driving force for densification, allowing dense nanocomposites to be manufactured at lower temperatures (e.g., 1500 °C for 5 vol.% SiC) and with higher volume fractions of second phase (e.g., 20 vol.% SiC at a temperature of 1675 °C) (Stearns and Harmer 1996).

The processing steps used to produce nanocomposites have a significant effect on their microstructures and properties. It is well known that the properties of the nanocomposite are strongly dependent on the second-phase dispersion (Stearns *et al.* 1992). Specifically, a uniform dispersion of second phase is critical to obtaining superior properties in the nanocomposite relative to the monolith. Unfortunately, the spatial distribution of the second phase has been shown to be very sensitive to the processing conditions used, indicating that small changes in processing can lead to significant differences in the properties of the nanocomposite.

The microstructures of nanocomposites produced using the above technique are also strongly dependent on the amount of second phase present. In nanocomposites with a high volume fraction of second phase (>5 vol.%), matrix grains are equiaxed with a narrow size distribution, and second-phase particles are uniformly distributed among the matrix grains. An example of this type of microstructure (5 vol.% SiC in Al_2O_3) can be seen in Fig. 1. On the contrary, in samples with a low volume fraction of second phase, the grain-size distribution of matrix grains may be significantly wider and in extreme cases, abnormal grains may be present.

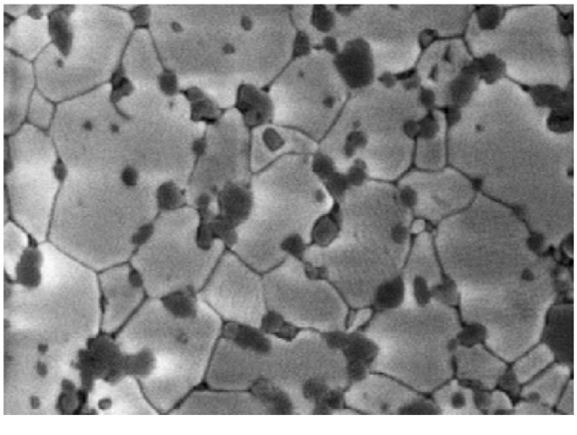


Figure 1
Scanning-electron-microscope image of the microstructure of a Al_2O_3–5 vol.% SiC nanocomposite showing the presence of both intergranular and intragranular particles.

---

* Cross references marked by an asterisk are included in this volume.

# Cross references marked by a hash can be found by consulting the Encyclopedia of Materials: Science and Technology.

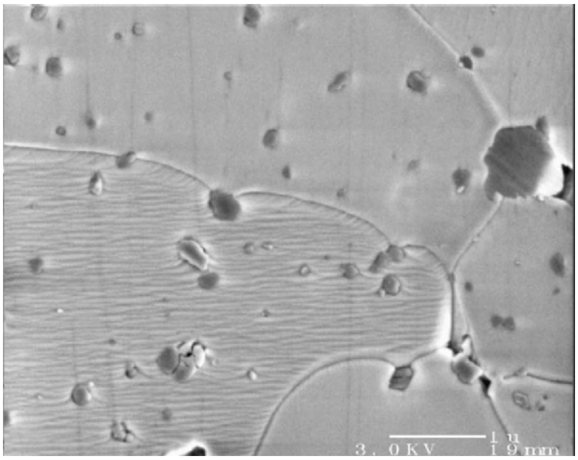


Figure 2
Scanning-electron-microscope image of SiC particles (5 vol.%) inhibiting grain growth by pinning migrating Al_2O_3 grain boundaries.

As shown in Fig. 1, second-phase particles in the nanocomposite are either located on grain boundaries (intergranular) or inside grains (intragranular). As will be discussed later, the location of particles with respect to grain boundaries plays the paramount role in the improvement of the properties of the nanocomposite over the monolith (Xu *et al.* 1997). In particular, intragranular particles are believed to be responsible for the superior room-temperature properties of the nanocomposite, while the intergranular particles are believed to be responsible for its superior high-temperature properties. Intergranular particles also significantly increase the microstructural stability of the nanocomposite. The interaction of particles with grain boundaries strongly inhibits matrix grain growth which would normally take place at high temperatures. Figure 2 shows SiC particles inhibiting grain growth in the nanocomposite by pinning migrating grain boundaries. The relative proportions of intra- and intergranular particles is strongly related to the matrix grain size in the Al_2O_3–SiC system, with the percentage of intergranular particles decreasing significantly with the grain size. For example, approximately 70% of particles are expected to be intergranular for a grain size of 1 μm, whereas 25% of particles are expected to be intergranular for a grain size of 5 μm (Stearns and Harmer 1996).

2. *Room-temperature Properties of Al_2O_3–SiC Nanocomposites*

The superior mechanical properties of Al_2O_3–SiC nanocomposites at room temperature were first shown by Niihara, who found that the addition of 5 vol.% of 300 nm SiC particles to Al_2O_3 increased the strength from 380 MPa to 1100 MPa (Niihara and Nakahira 1988). Furthermore, it was shown that a simple annealing treatment at 1300 °C for 1 hour in argon gas subsequently increased the strength of the nanocomposites to over 1500 MPa. Niihara reported that the addition of the SiC particles also resulted in a significant increase in the fracture toughness, from 3.8 $MPam^{1/2}$ to 4.7 $MPam^{1/2}$. The superior properties of Al_2O_3–SiC nanocomposites over monolithic Al_2O_3 have been confirmed by a number of researchers, although none of these improvements have been as dramatic as those reported by Niihara (Zhao *et al.* 1993, Borsa *et al.* 1995, Jang *et al.* 1996).

While the superior properties of the nanocomposite are well documented, the mechanisms responsible for the improvement in properties are still unclear. Proposed strengthening and toughening mechanisms for the unannealed nanocomposite include the reduction in critical flaw size due to the development of intragranular particle-induced subgrain boundaries, the attraction of intergranular cracks into matrix grains by residual tensile hoop stresses which arise due to the large thermal-expansion mismatch between Al_2O_3 and SiC, and the formation of a surface layer of compressive stresses during machining. The subsequent increase in strength and toughness after annealing has been attributed to the healing of surface flaws and the retainment of machining-induced compressive surface stresses, both the result of a slower stress relaxation behavior in the nanocomposite (Thompson *et al.* 1995).

In addition to having improved strength and toughness relative to monolithic Al_2O_3, Al_2O_3–SiC nanocomposites also have superior wear properties. The improved wear properties are the result of a change in fracture mode which occurs when the concentration of the SiC second phase is greater than ~4 vol.%. While monolithic Al_2O_3 exhibits predominantly intergranular fracture, the large thermal-expansion mismatch between Al_2O_3 and SiC results in an apparent strengthening of Al_2O_3 grain boundaries, causing the nanocomposite to fracture transgranularly. As a result, grains in Al_2O_3–SiC nanocomposites are worn away slowly, in contrast to monolithic Al_2O_3 where wear typically involves the removal of entire grains due to the relatively weak interface at grain boundaries. The effect of the change in fracture mode on the wear properties of the nanocomposite is dramatic. In one study, the wet erosive wear rate of a 5 vol.% SiC nanocomposite was measured to be one-third that of monolithic Al_2O_3 with the same grain size (Walker *et al.* 1994). Furthermore, the lack of grain pullout in the nanocomposite allows a higher-quality surface finish to be prepared (and retained during wear) than is possible with monolithic Al_2O_3.

3. High-temperature Properties of Al_2O_3–SiC Nanocomposites

In addition to its superior room-temperature properties, Al_2O_3–SiC nanocomposite also possesses superior high-temperature properties. The superior strength and toughness of the nanocomposite at room temperature are retained as high as 1000 °C (Niihara and Nakahira 1988). Furthermore, studies have shown that at higher temperatures, the creep resistance of Al_2O_3 is increased significantly by the presence of SiC particles (Ohji *et al.* 1995, Thompson *et al.* 1997). As shown in Fig. 3, the creep rate of Al_2O_3 containing 5 vol.% SiC (150 nm) is 2–3 orders of magnitude lower than that of monolithic Al_2O_3 at temperatures as high as 1300 °C. In addition, the creep life of the nanocomposite is more than an order of magnitude longer than in the monolith. Unfortunately, the failure strain of the nanocomposite is lower than that of Al_2O_3 due to the nucleation of cavities at intergranular particles. The improved creep resistance of Al_2O_3–SiC has been attributed to a number of mechanisms. These mechanisms include a reduction in the rate of grain-boundary diffusion due to the presence of intergranular particles; an inhibition of shape-accommodation processes necessary for creep to occur, such as grain-boundary sliding and migration; and a reduction of the intrinsic grain-boundary diffusivity of Al_2O_3 by silicon cation contamination. The dominant mechanism responsible for the improved creep behavior of the nanocomposite is unknown.

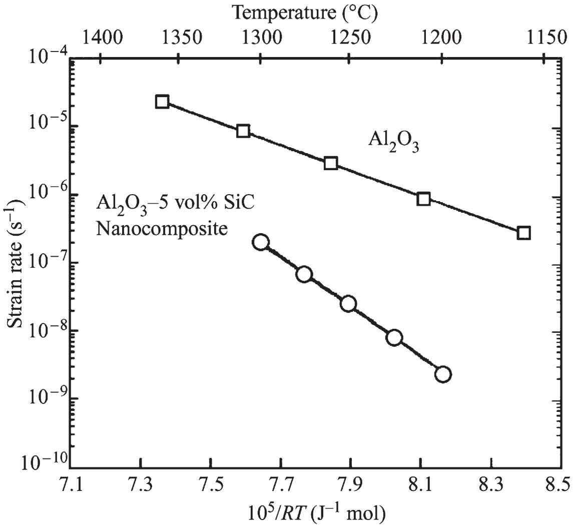


Figure 3
Plot of creep strain rate as a function of temperature for monolithic Al_2O_3 and Al_2O_3–5 vol.% SiC nanocomposites (after Thompson *et al.* 1997).

4. Concluding Remarks

Ceramic nanocomposites are relatively new materials which appear to have promise. In particular, Al_2O_3 containing nano-sized SiC second-phase particles has been shown to possess both room- and high-temperature properties which are superior to monolithic Al_2O_3. These properties, which include higher strength and toughness, increased wear resistance, lower creep rates at elevated temperatures, and excellent microstructural stability, make nanocomposites an attractive alternative to monolithic ceramics for many structural applications.

Bibliography

Borsa C E, Jiao S, Todd R I, Brook R J 1995 Processing and properties of Al_2O_3/SiC nanocomposites. *J. Microsc.* **177** (3), 305–12

Ding Z, Oberacker R, Thümmler F 1993 Microstructure and mechanical properties of yttria-stabilized tetragonal zirconia polycrystals (Y-TZP) containing dispersed silicon carbide particles. *J. Eur. Ceram. Soc.* **12** (5), 377–83

Jang B-K, Enoki M, Kishi T, Lee S-H, Oh H-K 1996 Fracture behavior and toughening of alumina-based composites fabricated by microstructural control. In: Bradt R C, Munz D, Sakai M, Shevchenko V Ya (eds.) *Fracture Mechanics of Ceramics*. Plenum, New York, Vol. 12, pp. 371–82

Jiao S, Jenkins M L, Davidge R W 1997 Electron microscopy of crack/particle interactions in Al_2O_3/SiC nanocomposites. *J. Microsc.* **185** (2), 259–64

Niihara K 1991 New design concept of structural ceramics: ceramic nanocomposites. *The Centennial Memorial Issue of The Ceramic Society of Japan* **99** (10), 974–82

Niihara K, Nakahira A 1988 Strengthening of oxide ceramics by SiC and Si_3N_4 dispersions . *Proc. 3rd Int. Symp. Ceramic Materials and Components for Engines*. The American Ceramic Society, Westerville, OH, pp. 919–26

Niihara K, Nakahira A 1990 Particulate-strengthened oxide nanocomposites. In: Vincenzini P (ed.) *Advanced Structural Inorganic Composites*. Elsevier, London, pp. 637–64

Niihara K, Nakahira A, Sasaki G, Hirabayashi M 1989 Development of strong Al_2O_3–SiC composites. *Materials Research Society Symp. Proc. Int. Meeting on Advanced Materials*. Materials Research Society, Warrendale, PA, Vol. 4, pp. 129–134

Ohji T, Nakahira A, Hirano T, Niihara K 1995 Tensile creep behavior of alumina/silicon carbide nanocomposite. *J. Am. Ceram. Soc.* **77** (12), 3259–63

Pezzotti G, Sakai M 1994 Effect of a silicon carbide "nano-dispersion" on the mechanical properties of silicon nitride. *J. Am. Ceram. Soc.* **77** (11), 3039–41

Stearns L C, Harmer M P 1996 Particle-inhibited grain growth in Al_2O_3–SiC: I, experimental results. *J. Am. Ceram. Soc.* **79** (12), 3013–9

Stearns L C, Zhao J, Harmer M P 1992 Processing and microstructure development in Al_2O_3–SiC "nanocomposites." *J. Eur. Ceram. Soc.* **10** (6), 473–7

Thompson A M, Chan H M, Harmer M P 1995 Crack healing and stress relaxation in Al_2O_3–SiC "nanocomposites." *J. Am. Ceram. Soc.* **78** (3), 567–71

Thompson A M, Chan H M, Harmer M P 1997 Tensile creep of alumina–silicon carbide "nanocomposites." *J. Am. Ceram. Soc.* **80** (9), 2221–8

Todd R I 1995 Alumina/SiC nanocomposites: big things from small particles. *Br. Ceram. Proc.* **55**, 225–38

Walker C N, Borsa C E, Todd R I, Davidge R W, Brook R J 1994 Fabrication, characterisation and properties of alumina-matrix nanocomposites. *Br. Ceram. Proc.* **53**, 249–64

Xu Y, Zangvil A, Kerber A 1997 SiC nanoparticle-reinforced Al_2O_3 matrix composites: role of intra- and intergranular particles. *J. Eur. Ceram. Soc.* **17**, 921–8

Zhao J, Stearns L C, Harmer M P, Chan H M, Miller G A 1993 Mechanical behavior of alumina-silicon carbide "nanocomposites." *J. Am. Ceram. Soc.* **76** (2), 503–10

G. S. Thompson and M. P. Harmer
Lehigh University, Bethlehem, Pennsylvania, USA

Natural Cellulose Fibers and Membranes: Biosynthesis

Cellulose is the most abundant organic material in nature. It forms the supporting framework in the cell walls of all plants, it occurs in the tissue of certain animals, and it is deposited extracellularly by some bacteria. Cellulose is a linear homopolymer with 1,4-β-glycosidic units in the chains with no known branching or irregularity but of various chain lengths. The biosynthesis of most natural cellulose comprises three steps: synthesis of the cellulose chains, formation of "mini sheets" of parallel chains by hydrophobic bonding, and assembly of the sheets to microfibrils. In some cases the cellulose biosynthesis is incomplete: only sheets or chains are formed and no microfibrils.

1. Crystal Structure and Morphology of Natural Cellulose

The natural cellulose deposited in plant cell walls is predominantly in the form of thin microfibrils (5–10 nm wide) containing parallel chains crystallized in a lattice called Cellulose I (Meyer and Mark 1929). The most stable lattice thermodynamically is Cellulose II, which forms when cellulose is precipitated from solution or recovered after swelling in strong alkali or acid (Rånby 1952). The parallel chain orientation in the Cellulose I microfibrils has been proven by several independent methods (Kuga and Brown 1987). All chains in a microfibril have the same direction.

A minor part of the cellulose in plant cell walls, e.g., in cotton fibers, is found to have Cellulose II lattice with antiparallel chains and occurring together with Cellulose I. Large amounts of Cellulose II is synthesized by mutated bacterial cells of Acetobacter in aberrant cultures and is shown to have folded chains (Kuga *et al.* 1993).

It was discovered by Atalla and Vanderhart (1984), using solid state NMR, that Cellulose I has two allomorphs, 1α and 1β, of which the 1β form is more stable than the 1α form. The two allomorphs occur together in the Cellulose I microfibrils in native plant cellulose. The animal cellulose in Tunicates (mantle animals) is synthesized in almost pure 1β form (Kimura and Itoh 1996).

The cell walls of certain algae, e.g., *Halicystis* spp., contain cellulose in the form of membranes without microfibrillar structure and with the thermodynamically stable Cellulose II lattice. These findings led to the tentative conclusion that the formation of cellulose microfibrils with the metastable Cellulose I lattice is an enzymatic process and not a crystallization of free cellulose chains (Rånby 1952).

The biosynthesis of cellulose is not a simple and straightforward process. It is a complex process with several different enzyme systems involved. A recent review describes the main steps and the known enzymes involved (Brown 1999).

2. Synthesis of Cellulose Chains

The monomeric glucose is formed from carbon dioxide and water by photosynthesis in green plants. By complex enzyme reactions, according to Leloir (1964) (Fig. 1), glucose-1-phosphate (1) is formed. It reacts with uridine triphosphate (2) to uridine diphosphate glucose (UDP-D-glucose) (3) and pyrophosphate (4). The UDP-D-glucose is polymerized by enzymes to cellulose chains, presumably by a stepwise reaction with release of UDP. There is a set of enzymes in plant cells simultaneously giving (1,3) glycosidic chains (callose) and (1,4)-β-chains (cellulose). Kudlicka and Brown (1997) have isolated and identified 12 polypeptides from mung beans (*Vigna radiate*) which are related to the formation of (1,4)-β-glucan chains in the cell walls. With these polypeptides present and separated from five other polypeptides forming (1,3)-glucans, pure (1,4)-β-glucan chains were formed and identified. With Acetobacter cells, corresponding enzymatic synthesis of cellulose chains has been verified.

There is evidence of a different pathway to UDP-D-glucose: sucrose and a sucrose synthase (SuSy) in plant cells react and form UDP-D-glucose (and then cellulose) and fructose according to Amor *et al.* (1995). This direct reaction would explain the fast and efficient cellulose synthesis in the heavy secondary wall of plant fibers.

3. Assembly to Cellulose Microfibrils

The discovery of the enzyme system for cellulose microfibril synthesis is a long story; Preston reported in

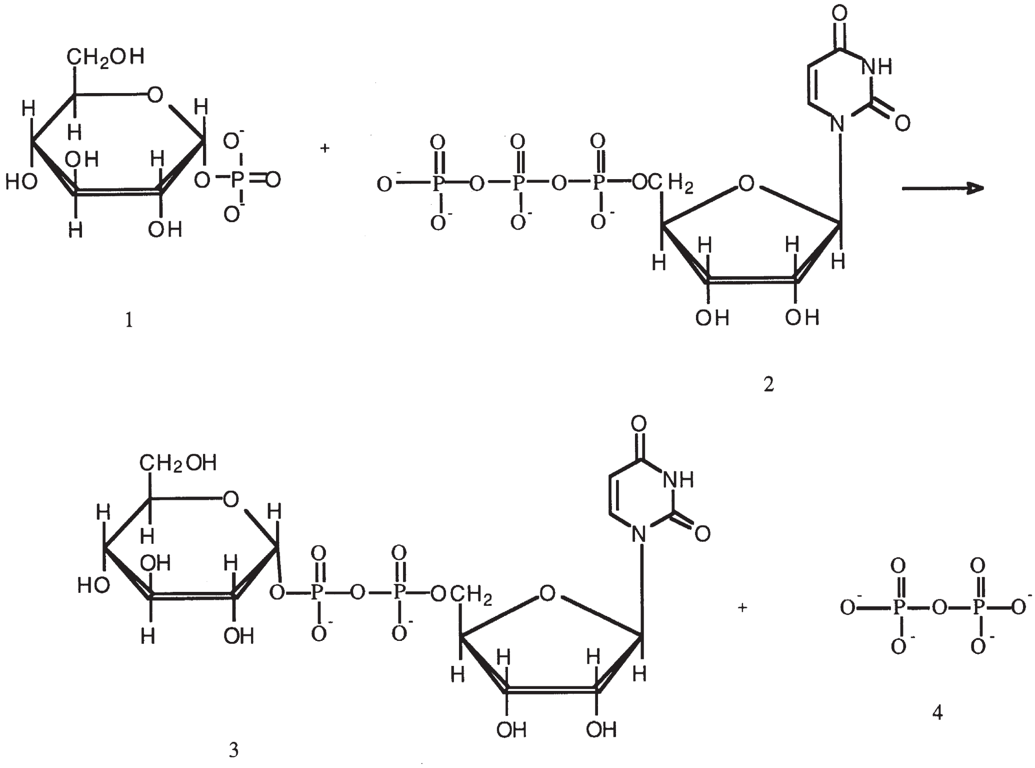


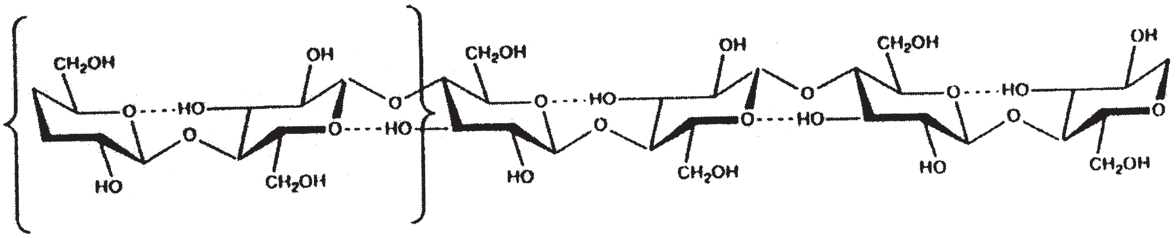
Figure 1
Enzymatic formation of the cellulose monomer uridine diphosphate glucose (UDP-D-glucose) (3) from a-D-glucopyranosyl-1-phosphate (1) and uridine triphosphate (2) under simultaneous release of pyrophosphate (4) according to Leloir (1964).

1964 that the cellulose microfibrils in wood fibers are associated with protein particles called "ordered granules" occurring together at the microfibril ends on the cell membrane and assumed to be the enzymes for cellulose synthesis.

In a study by Brown and Montezinos (1976) of the terminal complex (TC) particles, electron microscopic evidence was reported that the TC particles in the *Oocytis* algae were ordered-protein complexes of a linear type. They contain multiparticle protein subunits, associated with the tip of the growing microfibrils. In other organisms studied (mosses, ferns, and green algae) TC particles of both linear and rosette-like structure have been found to be associated with the growing cellulose microfibrils on the surface of the cell membrane (Fig. 2). Isolated intact rosette-like TCs in contact with UDP-glucose are shown to synthesize cellulose microfibrils *in vitro* (Kudlicka and Brown 1997), which strongly supports the view that the TC particles are the cellulose-forming enzymes which assemble microfibrils.

The dimensions of the microfibrils are related to the structure of the rosette-like TCs and their number of subunits. Valonia algae have large TSs with about 10 catalytic sites (=subunits per particle), which synthesize microfibrils with a thickness of about 30 cellulose chains, i.e., about 10 nm. *Vaucheria* algae have only one catalytic site per TC, which produces very thin microfibrils.

For cellulose chains synthesized enzymatically, Cousins and Brown (1995), have reported a second step where the parallel chains form "mini sheets" by hydrophobic bonding. The hydroxyl groups of the cellulose chains are on the sheet surfaces (Fig. 3). The parallel chains are held together in sheets by hydrophobic van der Waals forces. In a third step, the sheets are stacked together by hydrogen bonds and form elementary microfibrils with the Cellulose I lattice (Fig. 4). The specific stacking reaction appears to occur when the cellulose sheets (monomolecular layers) react with a connected complex enzyme system on the cytoplasmic membrane and form elementary cellulose microfibrils. This mechanism has been supported by the observation that the assembly of cellulose chains to microfibrils is prevented by the addition of a hydrogen-bonding compound, Tinopal,



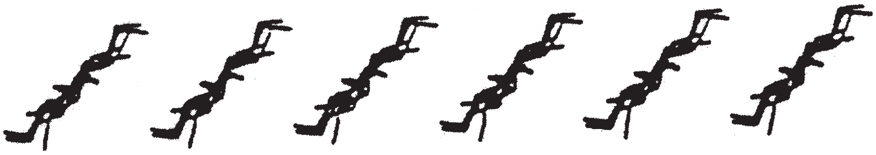


Figure 2
Schematic representation of formation of cellulose microfibril by a rosette terminal complex (TC) in plants (ferns, mosses, gymnosperms, and angiosperms).

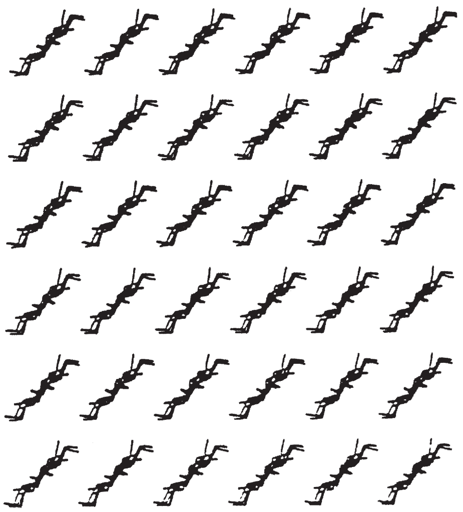


Figure 3
Cellulose chain (a) and monomolecular sheets of chains (mini sheets) held together with van der Waals forces (b) (Brown 1996).

LPW, which absorbs onto the cellulose chains and the "mini sheets" (Cousins and Brown 1997).

It was reported that plant cells of a mutant of Arabidopsis with incomplete rosette-like TC particles on the cell membrane produced single glucan chains, which fold and form Cellulose II lattices (Arioli *et al.* 1998).

4. Sequencing and Cloning of Cellulose-forming Enzymes

The first enzyme in the Cellulose synthase to be identified, sequenced, and cloned to a gene was a polypeptide (83 kDa) from *Acetobacter* (Saxena *et al.* 1990; Saxena *et al.* 1991). The active polypeptide had been purified before and applied for synthesis of 1,4-β-glucans (Lin *et al.* 1990). In later work, the enzymes and genes responsible for the biosynthesis of cellulose in *Acetobacter* and several eukaryotic organisms have been thoroughly studied (Saxena *et al.* 1994; Saxena and Brown 1995; Saxena *et al.* 1995).

Acetobacter forms cellulose microfibrils deposited by extracellular extrusion (Figs. 5 and 6). The active genes have been cloned. Corresponding genes have been found in cotton and rice (Delmer and Amor 1995) and defined as a Cel A homolog for cellulose synthase.

A tentative picture by Saxena (Fig. 5) of a cellulose-synthesizing enzyme complex in *Acetobacter xylinum* gives the main steps in microfibrillar formation (Brown 1996). The monomer UDP-D-glucose is synthesized by ACS AB gene products (= enzymes), the 1,4-β glucan chains by the sites indicated, "mini sheets" of the cellulose chain by the ACS D gene products and the stacking of the sheets to Cellulose I microfibrils by the ACS C gene products. For a mutated *Acetobacter* cell membrane (Fig. 6), with the ACS D gene inactive, the result is separate cellulose chains which are free to fold and give the observed stable Cellulose II lattice (Saxena *et al.* 1994). There is good evidence that the biosynthesis of cellulose in higher plant fibers, which occurs inside the cell wall,

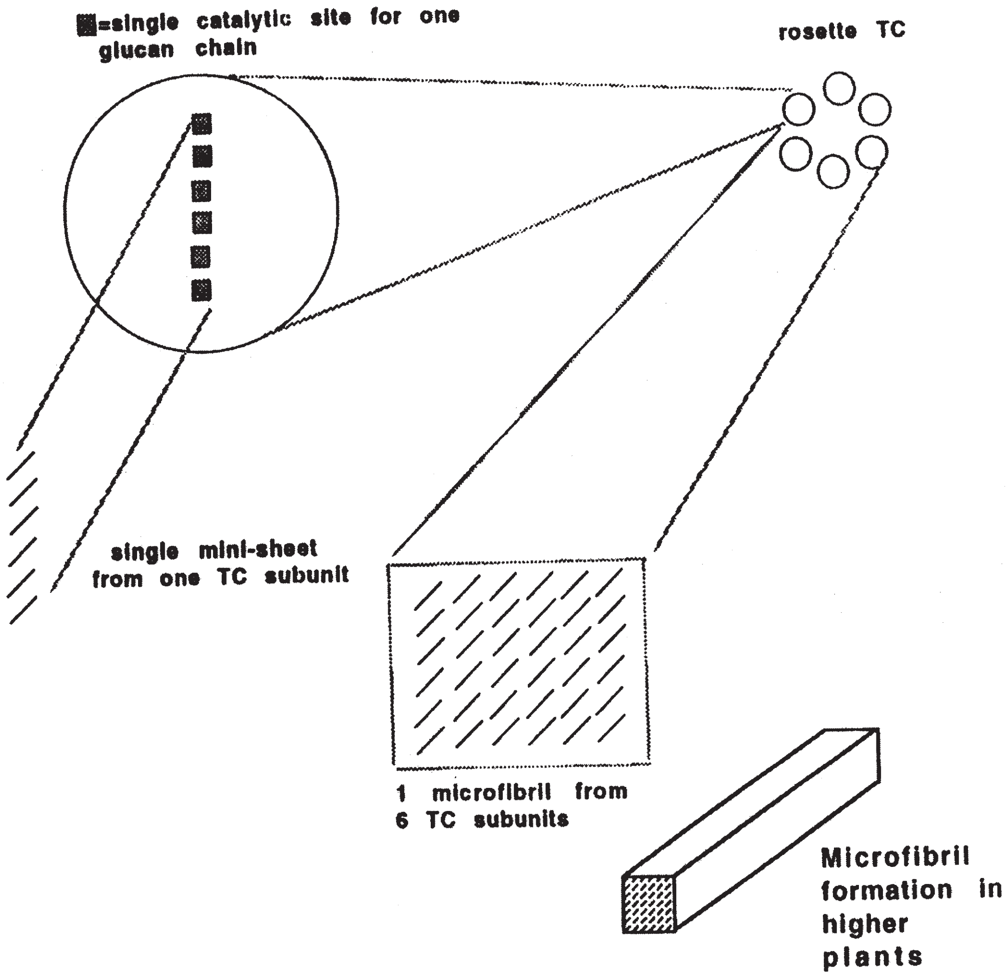

■=single catalytic site for one glucan chain
rosette TC
single mini-sheet from one TC subunit
1 microfibril from 6 TC subunits
Microfibril formation in higher plants

Figure 4
Cellulose microfibril formed by stacking mini sheets of cellulose chains to Cellulose I α lattice (Brown 1996).

is due to the same enzyme system as in *Acetobacter*. Cellulose and various polysaccharides are formed simultaneously in these fibers.

5. *Chemical* in vitro *Synthesis of Cellulose*

Cellulose chains and cellulose microfibrils with Cellulose I lattice have been synthesized *in vitro* by cellulose enzymes. Kobayashi *et al.* (1991) used β-cellobiosyl fluoride and an extracellular cellulose enzyme in a 5:1 acrylonitrile/acetate buffer and obtained cellulose chains which crystallized in the stable Cellulose II lattice (Fig. 7).

In later experiments, Kobayashi *et al.* (1994), in cooperation with Malcolm Brown, used more crude cellulose enzymes in a 2:1 acrylonitrile/acetate buffer and obtained microfibrillar cellulose with the metastable Cellulose I lattice. These are the first successful attempts of extracellular *in vitro* cellulose synthesis reported. The key is to use the enzymes under optimal conditions.

The first successful enzyme-free synthesis of cellulose chains has been reported by (Nakatsubo *et al.* 1996). They used a cationic ring-opening reaction, e.g., of glucose 2-orthopivaoyl 3,6-di-O-benzyl ether dissolved in the N,N′ carbonyldiimidazol of benzene under reflux, followed by treatment in Ph_3CBF_4/CH_2Cl_2 at room temperature. This synthesis has nothing to do with the biosynthesis but it shows an example of interesting possibilities for synthesizing specific polysaccharides and their derivatives using cationic ring-opening reactions.

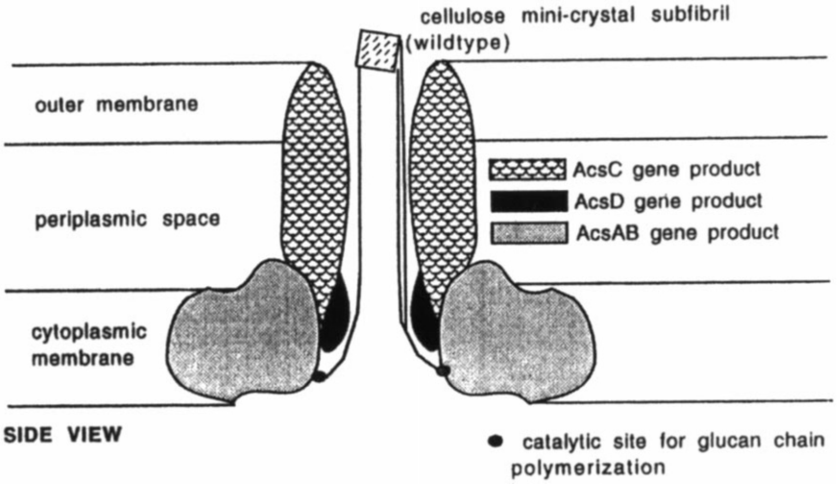

cellulose mini-crystal subfibril (wildtype)
outer membrane
periplasmic space
cytoplasmic membrane
SIDE VIEW
AcsC gene product
AcsD gene product
AcsAB gene product
● catalytic site for glucan chain polymerization

Figure 5
Schematic drawing of extracellular formation of a cellulose microfibril (Cellulose I) by the synthesizing complex of *Acetobacter*.

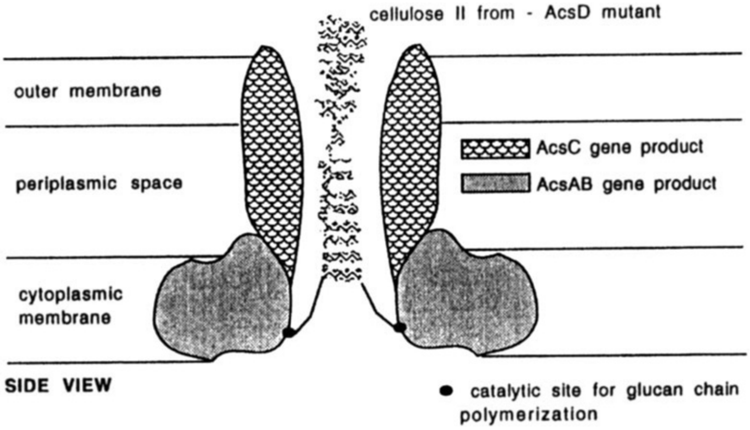

cellulose II from - AcsD mutant
outer membrane
periplasmic space
cytoplasmic membrane
SIDE VIEW
AcsC gene product
AcsAB gene product
● catalytic site for glucan chain polymerization

Figure 6
Corresponding synthesizing complex of a mutant which does not produce the ACS C gene product. The cellulose chains do not form mini sheets and are, therefore, free to fold and crystallize in Cellulose II lattice with antiparallel chains (Brown 1996).

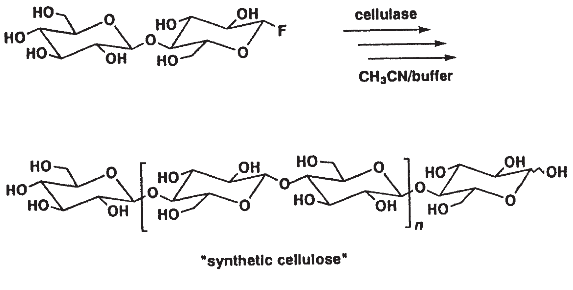

cellulase
CH₃CN/buffer
"synthetic cellulose"

Figure 7
Synthesis of cellulose chains *in vitro* using cellulose enzymes (Kobayashi *et al.* 1996).

6. *Conclusions*

The crystal structure and the morphology are outlined, describing the predominant microfibrils with Cellulose I lattice and the cellulose membranes crystallizing with the thermodynamically more stable Cellulose II lattice. The biosynthesis of cellulose is described as a three-step process: the synthesis of cellulose chains from the monomer uridine diphosphate glucose (UDP-D-glucose), the formation of monomolecular "mini sheets," and the stacking of the mini sheets to cellulose microfibrils with Cellulose I lattice. The corresponding enzymes are isolated and cloned to genes, which form the cellulose synthase system in plants, ferns, and mosses, and a few bacteria and primitive animals. The enzyme system forms terminal complex (TC) particles from which the cellulose microfibrils are extruded. Certain algae, e.g., *Halicystis,* and mutated *Acetobacter*, have an incomplete cellulose synthase system and form cellulose chains that are folded and crystallize in the more stable Cellulose II lattice.

Acknowledgment

Minor final editing by Paul Calvert, following the decease of the author in 2000.

See also: Textile Fibers: A Comparative Overview*; Natural Cellulose Fibers: Properties*

Bibliography

Amor Y, Haigler C, Johnson S, Wainscote M, Delmer D P 1995 A membrane-associated form of sucrose synthase and its potential role in synthesis of cellulose and callose in plants. *Proc. Natl. Acad. Sci. USA* **92**, 9353–7

Arioli T, Peng L, Betzner A S, Burn J, Wittke W, Herth W, Camilleri C, Hoffe H, Plazinski J, Birch R 1998 Molecular analysis of cellulose biosynthesis in Arabidopsis. *Science* **279**, 717–20

Atalla R, Vanderhart D 1984 Native cellulose—a composite of two distinct crystalline forms. *Science* **223**, 283–5

Brown R 1996 The biosynthesis of cellulose. *J. Macromol. Sci. A* **33**, 1345–73

Brown R 1999 Cellulose structure and biosynthesis. *Pure Appl. Chem.* **71**, 767–75

Brown R, Montezinos D 1976 Cellulose microfibrils—visualization of biosynthetic and orienting complexes in association with plasma membrane. *Proc. Natl. Acad. Sci. USA* **73**, 143–7

Cousins S, Brown R 1995 Cellulose-I microfibril assembly—computational molecular mechanics energy analysis favors bonding by van der Waals forces as the initial step in crystallization. *Polymer* **36**, 3885–8

Cousins S, Brown R 1997 X-ray diffraction and ultrastructural analyses of dye-altered celluloses support van der Waals forces as the initial step in cellulose crystallization. *Polymer* **38**, 897–902

Delmer D, Amor Y 1995 Cellulose biosynthesis. *Plant Cell* **7**, 987–1000

Kimura S, Itoh T 1996 New cellulose synthesizing complexes (terminal complexes) involved in animal cellulose biosynthesis in the tunicate. *Metandrocarpa uedai. Protoplasma* **194**, 151–63

Kobayashi S, Kashiwa K, Kawaski T, Shoda S 1991 Novel method for polysaccharide synthesis using an enzyme—the first *in vitro* synthesis of cellulose via a nonbiosynthetic path utilizing cellulase as catalyst. *J. Am. Chem. Soc.* **113**, 3079–84

Kobayashi S, Okamoto E, Wen X, Shoda S 1996 Chemical synthesis of native-type cellulose and its analogues via enzymatic polymerization. *J. Macromol. Sci. A* **33**, 1375–84

Kobayashi S, Shoda S, Lee J, Okuda K, Brown R M 1994 Direct visualization of synthetic cellulose formation via enzymatic polymerization using transmission electron-microscopy. *Macromol. Chem. Phys.* **195**, 1319–26

Kudlicka K, Brown R 1997 Cellulose and callose biosynthesis in higher plants.1. Solubilization and separation of (1->3)- and (1->4)-beta-glucan synthase activities from mung bean. *Plant Physiol.* **115**, 643–56

Kuga S, Brown R 1987 Lattice imaging of ramie cellulose. *Polym. Commun.* **28**, 311–4

Kuga S, Takagi S, Brown R M 1993 Native folded chain cellulose-II. *Polymer* **34**, 3293–7

Leloir L 1964 Nucleoside diphosphate sugars + saccharide synthesis. *Biochem. J* **91**, 1

Lin F C, Brown R M, Drake R R, Haley B E 1990 Identification of the uridine 5′-diphosphoglucose (UDP-GLC) binding subunit of cellulose synthase in acetobacter-xylinum using the photoaffinity probe 5-azido-UDP-GLC. *J. Biol. Chem.* **265**, 4782–4

Meyer K, Mark H 1929 *Z. Phys. Chem.* **B2**, 115

Nakatsubo F, Kamitakahara H, Hori M 1996 Cationic ring-opening polymerization of 3,6-di-O-benzyl-alpha-D-glucose 1,2,4-orthopivalate and the first chemical synthesis of cellulose. *J. Am. Chem. Soc.* **118**, 1677–81

Rånby B G 1952 *Acta. Chem. Scand.* **6**, 101–15, 116–27, 128–40

Saxena I, Brown R 1995 Identification of a second cellulose synthase gene (Acsaii) in *Acetobacter xylinum. J. Bacteriol.* **177**, 5276–83

Saxena I M, Brown R M, Fevre M, Geremia R A, Henrissa B 1995 Multidomain architecture of beta-glycosyl transferases—implications for mechanism of action. *J. Bacteriol.* **177**, 1419–24

Saxena I M, Kudlicka K, Okuda K, Brown R M 1994 Characterization of genes in the cellulose-synthesizing operon (acs operon) of *Acetobacter xylinum*—implications for cellulose crystallization. *J. Bacteriol.* **176**, 5735–52

Saxena I M, Lin F C, Brown F M 1990 Cloning and sequencing of the cellulose synthase catalytic subunit gene of *Acetobacter xylinum. Plant Mol. Biol* **15**, 673–83

Saxena I M, Lin F C, Brown F M 1991 Identification of a new gene in an operon for cellulose biosynthesis in *Acetobacter xylinum. Plant Mol. Biol.* **16**, 947–54

B. Rånby[†]
Royal Institute of Technology, Stockholm, Sweden

Natural Cellulose Fibers: Properties

Cellulose is the main component of the plant cell wall, with a share of 60–90%. At the same time, it is the main strengthening element of this natural composite material. The mechanical properties are quite remarkable: a cellulose filament of 1 mm diameter can hold more than 60 kg in weight—it has 80% of the strength of steel. Mechanical properties for a number of different fibers are given in Table 1.

The cellulose is incorporated into the cell wall in form of thin fibrils, so called *microfibrils* or *elementary fibrils*, with an elaborate substructure (Fig. 1). Their diameter measures 35 Å. The native cellulose is an unbranched polymer of β-1-4-linked glucose residues arranged in linear, antiparallel chains, where every other glucose residue is rotated at approximately 180°. As a result, cellobiose is the structural

Table 1
Mechanical properties of different fibers.

Fiber	Wet or dry	Tensile modulus (GPa)	Tensile strength (MPa)	Extension to break (%)
Hemp	Wet	35		
	Dry	70	920	1.7
Jute	Wet			3
	Dry	60	860	
Flax	Wet	27	880	2.2
	Dry	95	840	1.8
Ramie	Wet	19	1080	2.4
	Dry	80	920	2.3
Cotton	Wet	7.5	240–830	
	Dry	27	200–800	6–12
Collagen	Wet	2	100	8–10
Silk		0.3	600	18–20
Rayon	Wet	0.4	130	28
	Dry		260	27
Fortisan, high strength, regenerated cellulose	Wet	4.4	650	8
	Dry		900	6
Lyocell	Wet		580	
	Dry	4.3	650	15
Polyester	Dry	7	720	44
	Wet		720	

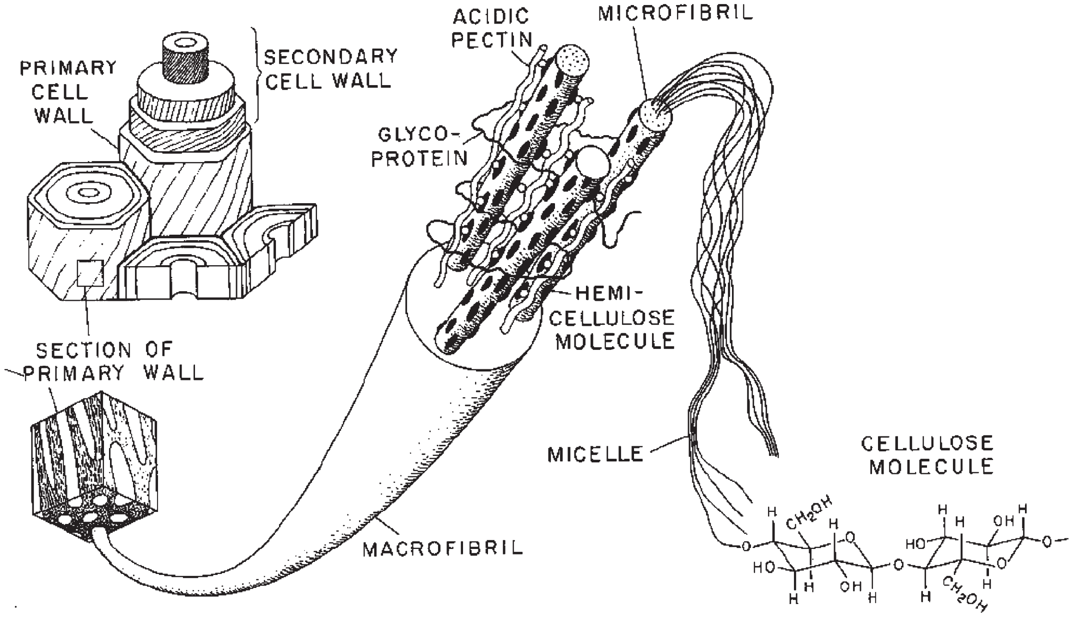

PRIMARY CELL WALL
SECONDARY CELL WALL
ACIDIC PECTIN
MICROFIBRIL
GLYCO-PROTEIN
HEMI-CELLULOSE MOLECULE
SECTION OF PRIMARY WALL
MICELLE
CELLULOSE MOLECULE
MACROFIBRIL
CH2OH
HO
OH
H
O

Figure 1
Overview over the various levels of organization within plant cell walls (after Niklas 1984).

repeating unit of the glucan chains in cellulose. This is in contrast to other glucan polymers such as starch where glucose is the repeating unit. In the native cellulose microfibril crystalline regions alternate with unordered, amorphous regions. From x-ray diffraction and negative staining with subsequent electron microscopic observation the length of the crystalline regions has been calculated to be 60 nm. Crystalline cellulose can take on two different polymeric structures—the extended-chain configuration (cellulose I) and the folded-chain configuration (cellulose II, Fig. 2). While the former is found in native cellulose and is the thermodynamically more stable form, the latter arises by crystallization from solution, where it is unlikely that the random-coil molecules will extend fully to fit into an extended-chain crystal. The folded-chain configuration can thus be viewed as an energetic compromise. In the plant cell wall cellulose I can form because here the synthesis and crystallite formation take place simultaneously and there is no solution phase (see *Natural Cellulose Fibers and Membranes: Biosynthesis**). Regenerated, technical cellulose such as rayon always has the folded-chain configuration with lower values of modulus and tensile strength than native cellulose. Interestingly, cellulose II forms fibrils of the same diameter as cellulose I. This first led to the conclusion that they would be identical. This caused a lot of confusion as regards the crystal structure of native cellulose.

It is not possible to measure the mechanical properties of native cellulose directly. In the plant cell it is always found together with different matrix substances and upon isolating it loses its native conformation. Thus, theoretical approaches were made to calculate these properties from the molecular structure and the interatomic forces. The resulting values range from 60 GPa (Treloar 1960) to 320 GPa (Gillis 1969). Experimental measurements have been made on cells consisting almost exclusively of cellulose, e.g., ramie fibers. More sophisticated studies have derived the modulus of cellulose from x-ray measurements on strained ramie fiber, where the specific strain of the cellulose fibrils was calculated from changes in the x-ray diffraction angles (Sakurada *et al.* 1962). The experimental values for bast fibers go up to 110 GPa (Meyer and Lotmar 1936). The theoretical values should be higher, since they do not take into account the amorphous regions of the microfibril but only the much stiffer crystallites.

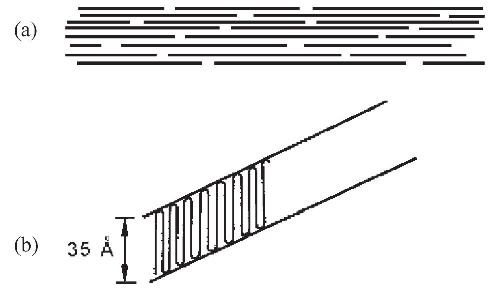


Figure 2
Polymer configurations: (a) extended-chain (b) folded-chain.

For theoretical calculations different models have been used concerning the accuracy of the employed molecular model and the interatomic forces taken into account. A very comprehensive model has been published by Gillis (1969) and is based on the real 3D anisotropic crystal lattice of cellulose in its native state. The molecular processes occurring upon straining are mainly interchain hydrogen bond deformations and alterations in the valence angle at the bridging oxygen atoms between the monomers. The necessity of these two processes can be derived from the schematic representation of a unit cell depicted in Fig. 3. Straining will align the individual molecular strands and necessarily deform the hydrogen bonds in angle and length together with a change in the valence angle of the bridging oxygen atom. In tension, deformation of the bridging oxygen bond plays the most important role, accounting for about 50% of the deformation. These oxygen bridges also constitute the limiting factor with regards to strength: the multiple interchain hydrogen bonding is so strong that the failure of cellulose fibrils is not due to slipping of adjacent chains but to chain scission.

In the plant cell wall the cellulose microfibrils are embedded in an aqueous medium, which affects their mechanical properties. The water penetrates the amorphous regions and competes for potential hydrogen bonding sites, loosening the interactions between

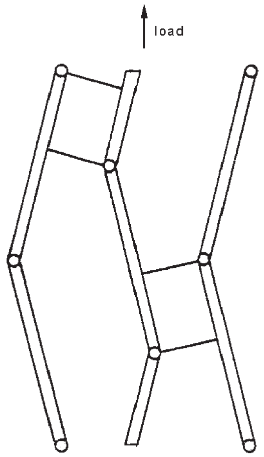


Figure 3
Crystalline unit cell of inextensible glucose residues linked in the chain direction by oxygen atoms and cross-linked by hydrogen bonds (after Gillis 1969).

adjacent crystalline regions. This effects an increase in both extensibility and strength of the fibrils, while the modulus of elasticity is decreased. A counteracting effect to the hydration is exerted by deposition of hydrophobic lignin in the cell wall, which waterproofs it. This opens the possibility of a specific adjustment of the mechanical properties. For bast fibers the difference between wet and dry in the modulus of elasticity amounts roughly to a factor of 3.5. In wood, which is heavily lignified, the factor is about half. A further modification of the cell walls' overall mechanical properties is made on the structural level through orientation of the microfibrils. In cells subjected mainly to tension, such as fiber cells, the fibrils are predominately oriented parallel to the longitudinal axis. Some cells, such as wood cells, also have to cope with tension, with bending, with negative pressure for water transport and with damping of vibration. Here the fiber axes are tilted for a trade-off between these different demands (Booker and Sell 1998).

Bibliography

Booker R E, Sell J 1998 The nanostructure of the cell wall of softwoods and its functions as a living tree. *Holz als Roh- und Werkstoff* **56**, 1–8

Gillis P P 1969 Effect of hydrogen bonds on the axial stiffness of crystalline cellulose. *J. Polym. Sci.* A **2**, 783–794

Meyer K H, Lotmar W 1936 Sur l'élasticité de la cellulose. *Helv. Chim. Acta* **19**, 68–86

Niklas K J 1984 *Plant Biomechanics*. University of Chicago Press, Chicago, IL

Sakurada I, Nukushina Y, Ito T 1962 Experimental determination of the elastic modulus of crystalline regions in oriented polymers. *J. Polym. Sci.* **57**, 651–60

Treloar L R G 1960 Calculations of elastic moduli of polymer crystals. III. Cellulose. *Polymer* **1**, 290–303

L. Köhler
University of Freiburg, Germany

Natural Protein Fibers

Many of Nature's functional materials are produced as fibers. This geometry allows the covalent bonds in polymer backbones to be aligned optimally along directions where reinforcement is needed. A fibrous morphology is also compatible with the short transport distances needed to expedite the self-assembly of materials from an aqueous environment. Larger cross-sections are self-assembled efficiently by lateral association of adjacent fibers, rather than depending on the growth of a thick monofilament.

Water plays a pivotal role in the biological self-assembly of fibers. It promotes liquid crystallinity and thus supramolecular self-assembly. It can help to maintain the structural and functional integrity of the assembled fibers (e.g., muscle, cytoskeleton, and virus spikes), and it provides a medium in which the biological mechanisms of assembly and self-repair can be effected.

Natural protein fibers include collagen, keratin, silk, mussel byssus, viral spikes, tubulin, and actin. All exhibit hierarchical molecular order. Their properties therefore depend on structure over a range of length scales as well as on the composition of the individual molecules. Hierarchical molecular order promotes both damage resistance and damage tolerance. It also enables a material to be multifunctional because different structural features, at different length scales, can be tailored to optimize different properties. The optimized properties need not all relate to load-bearing ability; they may include responsiveness to electrical stimuli (muscle fibers), to chemical information (insect antennae), or to ambient water (capture threads in spider webs).

1. Background

1.1 Technological Interest in Natural Protein Fibers

Fibrous proteins have been exploited directly for hundreds—in some cases thousands—of years. Examples include silk and wool harvested for textile applications, the use of intestinal collagen to stitch wounds or to "string" instruments of music and sport, and hair serving as the moisture sensor in hygrometers. However, it was late in the twentieth century before materials characterization technology became sufficiently advanced to reveal detailed links between structure and function in these complex materials.

Studies of natural fibers promise a number of potentially useful lessons for materials chemistry and processing. Nature's range of functional materials represents the success stories of up to four billion years of research and development. Nature has optimized supramolecular self-assembly mechanisms, hierarchical microstructures, property combinations, and in-service durability. Nature's fibrous materials are not only damage-tolerant but often self-repairing. They offer the attractions of biosynthesis (they are produced from renewable resources), benign processing conditions (they are assembled and shaped in an aqueous environment and at mild temperatures), and biodegradability (they break down into harmless components when exposed to some specific natural environments).

The techniques of molecular biology can be used to genetically engineer host cells or even multicellular organisms that are capable of synthesizing economic quantities of protein for possible processing into fiber. Proteins that already exist in Nature, as well as entirely new materials, can be produced in this way.

1.2 Technological Limitations of Proteins in Fiber Engineering

Many materials that are indispensable in technology have no counterpart in Nature because natural selection has exerted no pressure to evolve such materials. Thus, natural proteins cannot compete with carbon fiber in regard to high temperature applications, or with conjugated polymers in regard to electrical conductivity.

Natural mechanisms of materials production are usually complex and/or slow in comparison to conventional artificial methods. Spiders produce fiber more rapidly than any other organism, but the maximum rates are around $10\,\mathrm{cm\,s^{-1}}$ and are not maintained continuously. This maximum is two to three orders of magnitude slower than typical rates for industrial dry solution spinning. The need to achieve economical fiber production rates while applying Nature's lessons for "green" materials processing and microstructure–property optimization represents a formidable challenge.

Proteinaceous materials are sensitive to both moisture and temperature. In some cases, water may be essential for a protein fiber to exhibit optimum properties, e.g., the energy-dissipating ability of capture thread in a spider's web, or the compressive strength of viral coat protein spikes. Such fibers would be of little use under dry conditions.

There are many factors that limit the viability of genetically altered organisms as a source of functional polymers: the stability of the altered gene (to mutation as well as to repair), toxic effects of the foreign protein on the host organism, intracellular degradation of the foreign protein by the host's enzymes, and the extent to which the desired protein can be recovered in useful amounts and in a processable form.

1.3 Fiber Self-assembly

Successful fiber self-assembly requires molecules that can serve as appropriate building blocks, a source of structural anisotropy, an environment that promotes self-assembly, and intermolecular interactions (bonds) that stabilize the assembled fiber.

(*a*) *Proteins as versatile fiber-forming materials.* The natural monomers of proteins are α-amino acids –(CO–CHR–NH)$_n$–. Following protein synthesis, some of the R-groups (Fig. 1) may undergo derivatization, effectively increasing the diversity of monomer types in the polymer. The unique properties of a given protein depend on which amino acids are present, the relative amounts of those amino acids, the specific sequence in which they are arranged, and the overall length of the molecule. The variety of possible proteins is therefore extremely large, limited only by the practicalities (i.e., kinetic considerations) of getting long molecules to fold into a conformation that ensures the desired properties.

Two characteristics of protein molecules underlie their versatile fiber-forming ability:

(i) The backbones are intrinsically flexible, so proteins have the potential to fold into a large variety of possible conformations.

(ii) The R-groups (Fig. 1) confer a locally hydrophilic (acidic, basic, or neutral) or hydrophobic character to the molecules. In aqueous environments, the charge distribution along the molecules will therefore be sensitive to the presence of other charged species, and so will depend on pH and dissolved ions. The flexible protein chains will tend to fold in a way that allows hydrophobic segments to be screened from the water by the more hydrophilic segments, while also tending to maximize the extent to which charged sites participate in hydrogen bonding.

It follows that the three-dimensional structure of a protein depends on the nature and sequence of the amino acid side groups along the molecule, the related backbone flexibility, the characteristics of the aqueous environment, and the characteristics of neighboring protein molecules.

(*b*) *Proteins as a source of structural anisotropy.* The intrinsic flexibility of proteins allows them to form geometrically anisotropic structures such as helices or β-sheet crystals (Fig. 2). This occurs in water, and can be promoted by hydrogen bonding and hydrophobic interactions. Structural stability is sensitive to the amino acid sequence of the protein, and to the pattern of intramolecular and intermolecular bonding favored by interactions between the protein and its environment. Polymer–water interactions are particularly significant. Some fibrous proteins, such as those found in the byssal threads of mussels, exhibit a hybrid structure which incorporates collagen-like (helix-forming) and silk-like (sheet-forming) sequences to achieve a gradient in mechanical properties along the fiber.

Protein molecules which instead have a globular conformation can form anisotropic structures by aggregating (Fig. 2). This process can be promoted by an asymmetric charge distribution (dipoles) on the interacting macromolecules, and by specificity of sites that take part in intermolecular hydrogen and hydrophobic bonds. Ions may also play a role in binding.

The nucleation of anisotropic supramolecular aggregates from helices, β-sheets or globular molecules can be promoted by molecular templates. Examples include the extension peptides which aid the assembly of collagen triple helices, templates for the hollow triangular prisms (consisting of three cross-β-sheet crystals) that constitute viral spikes, and the microtubule organizing centers (MTOCs) that assist the formation of microtubules in the cytoskeleton and flagellae.

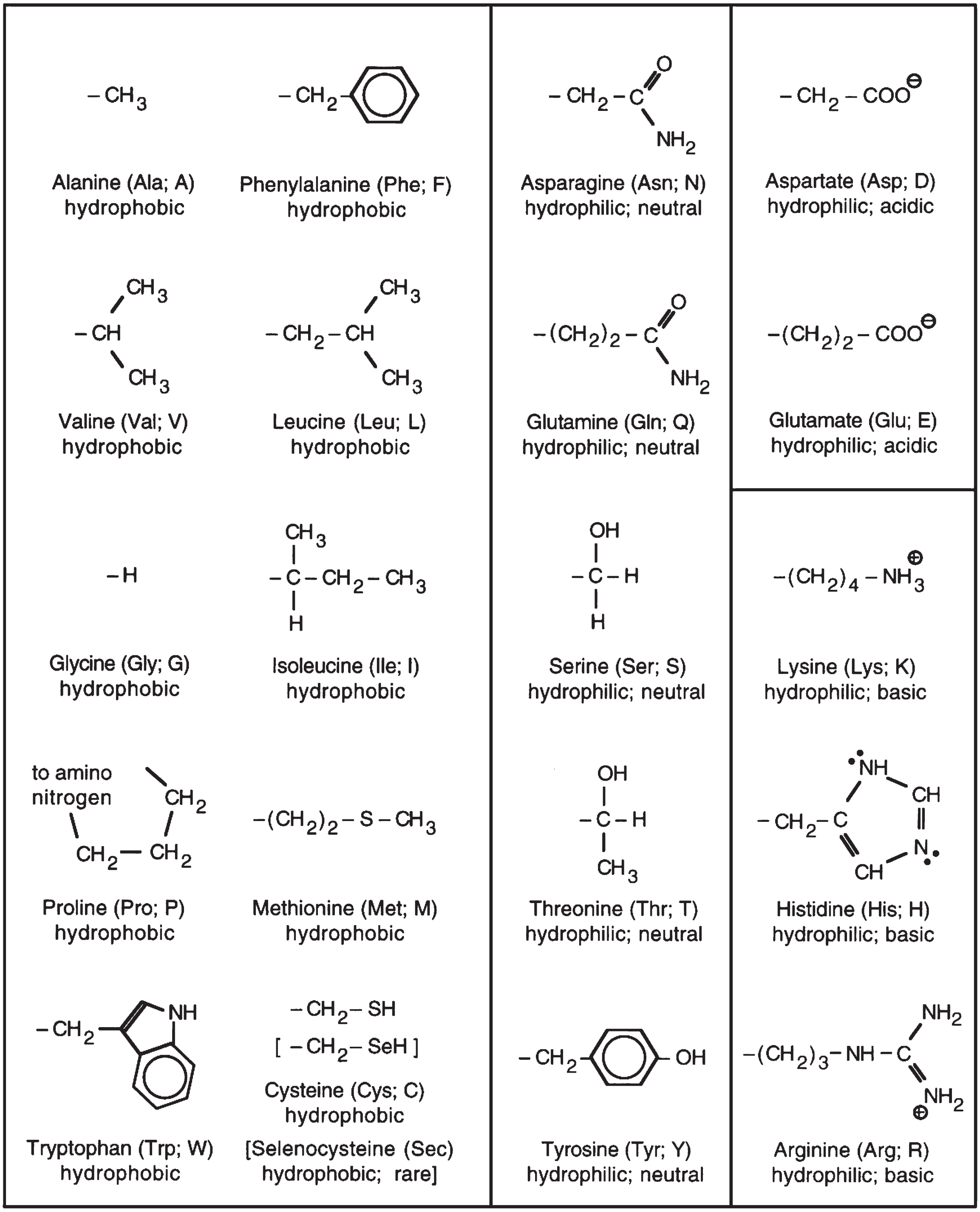

Alanine (Ala; A)
hydrophobic
Phenylalanine (Phe; F)
hydrophobic
Asparagine (Asn; N)
hydrophilic; neutral
Aspartate (Asp; D)
hydrophilic; acidic
Valine (Val; V)
hydrophobic
Leucine (Leu; L)
hydrophobic
Glutamine (Gln; Q)
hydrophilic; neutral
Glutamate (Glu; E)
hydrophilic; acidic
Glycine (Gly; G)
hydrophobic
Isoleucine (Ile; I)
hydrophobic
Serine (Ser; S)
hydrophilic; neutral
Lysine (Lys; K)
hydrophilic; basic
to amino nitrogen
Proline (Pro; P)
hydrophobic
Methionine (Met; M)
hydrophobic
Threonine (Thr; T)
hydrophilic; neutral
Histidine (His; H)
hydrophilic; basic
Tryptophan (Trp; W)
hydrophobic
Cysteine (Cys; C)
hydrophobic
[Selenocysteine (Sec)
hydrophobic; rare]
Tyrosine (Tyr; Y)
hydrophilic; neutral
Arginine (Arg; R)
hydrophilic; basic

Figure 1
The side groups that distinguish the 21 α-amino acids from which all known natural proteins are synthesized. Conventional three-letter and one-letter abbreviations for each amino acid are shown, along with the characteristics of the side groups in aqueous environments.

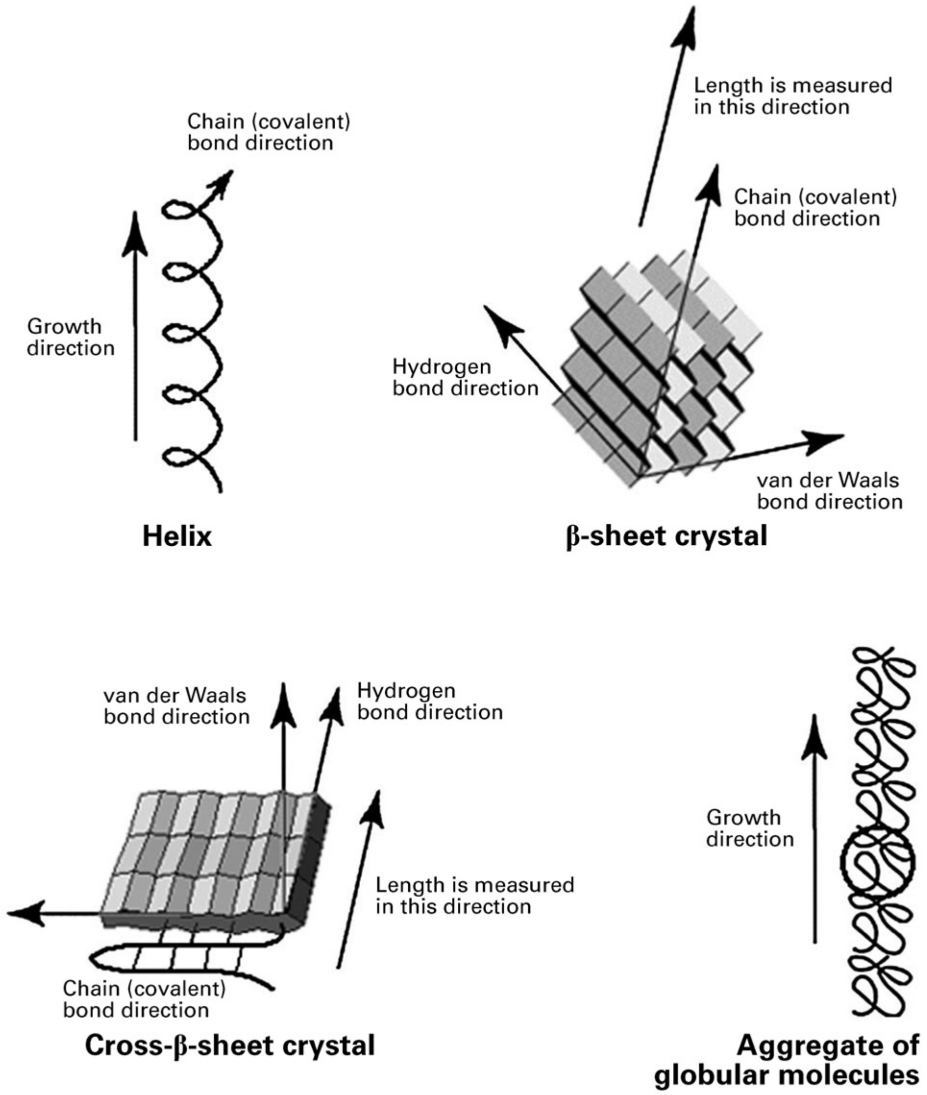

Chain (covalent)
bond direction
Growth
direction
Helix
Length is measured
in this direction
Chain (covalent)
bond direction
Hydrogen
bond direction
van der Waals
bond direction
β-sheet crystal
van der Waals
bond direction
Hydrogen
bond direction
Length is measured
in this direction
Chain (covalent)
bond direction
Cross-β-sheet crystal
Growth
direction
Aggregate of
globular molecules

Figure 2
Four commonly occurring anisotropic structures formed by protein molecules.

Liquid crystalline order can also promote the aggregation of globular molecules into anisotropic structures. In F-actin (muscles), in microtubules, and in silk secretions prior to spinning, synergism develops between alignment-enhanced assembly and shape-enhanced alignment.

(*c*) *Propagation of structural anisotropy: self-assembly at greater length scales.* In an aqueous environment, helices can form superhelices (coiled coils) stabilized by hydrophobic bonding across the area of helix-helix contact. Superhelix formation may alternatively be promoted by electrostatic interactions. Liquid crystalline order can directly control the higher level pattern into which rod-like structures become organized. There are many distinguishable schemes according to which rods can pack in a liquid crystalline fluid, and all the major classes have been detected in protein fibers.

2. Examples of Natural Fibrous Materials Based on Helices

2.1 Keratins

The characteristic α-helical conformation of keratins, first described by Pauling and Corey in 1951, is one of the simplest regular conformations that a sequence of amino acids in a protein can adopt. When it occurs in fiber-forming proteins, it tends to account for all the secondary structure, i.e., the entire molecule is α-helical. The structure of keratin is described in detail in *Keratin*[#].

2.2 Collagens

(*a*) *Collagen helices.* The individual linear unbranched polymer chains in collagen are called *procollagen.* Their general formula is (Gly-X–Y)$_n$, where many of the amino acid residues X and Y are proline. The inflexibility of the backbone nitrogen–carbon (N–C) bond within proline (Fig. 1) ensures the formation of helical conformations that are more stretched out than the α-helix in keratins. Many of the proline residues—and also some lysine residues—undergo modification (hydroxylation), after which they are able to form hydrogen bonds and covalent cross-links that stabilize the supramolecular structure of collagen.

The left-handed procollagen helices self-assemble into right-handed triple-helical supercoils. Assembly and stability are favored by several factors:

(i) The sequence of amino acid residues at one end of the procollagen molecule includes cysteine. The molecule is therefore capable of forming disulfide bonds with the corresponding sequence on an adjacent procollagen. (The cysteine-containing sequences are known as *extension peptides*, because they are discarded at a later stage of collagen synthesis.)

(ii) Intramolecular and intermolecular hydrogen bonding is extensive.

(iii) The amino acid sequence and the helical conformation of procollagen together generate a hydrophobic stripe of glycine residues along one side of the helix. Glycine has a small side group (hydrogen), so there is room to accommodate the glycine residues in the space between three parallel close-packed helices.

Attempts to produce artificial analogs of such a structure must therefore recognize that there are several superimposed contributions to stability in the natural material. Some of the polymer–polymer interactions involve the localization of water molecules to participate in hydrogen bonding, while other interactions specifically involve the local exclusion of water.

(*b*) *Tropocollagen and collagen.* Subsequent stages in the hierarchical assembly of collagen occur outside the cell. The extension peptides are cleaved off by enzymes, leaving *tropocollagen* triple helices. Tropocollagen then self-organizes into a wide variety of structures, many of which exhibit features typical of liquid crystalline order.

(*c*) *Properties.* Molecular and supercoil twists in tropocollagen occur in opposite directions. This rope-like structure ensures that tropocollagen is stiff and strong in tension. At coarser length scales there can be additional levels of "rope-within-rope" structural hierarchy, as well as pleated structures (e.g., in tendon) that enhance extensibility without sacrificing ultimate tensile strength. The hierarchical microstructure of collagen confers outstanding damage tolerance and toughness. The bulk properties of collagen depend on the orientation of fibers relative to the loading geometry, the degree and nature of cross-linking within fibers (which increases with age), moisture content, temperature, the relative amounts of fiber and any matrix phase, and the rheology of the matrix.

More than a quarter of the protein in mammals is collagen, much of it existing as fibers in tendons, ligaments, cartilage, bones, and teeth. The breaking strength of rat tail tendon is approximately 60 MPa, while the tensile modulus depends on the age of the rat. Collagen is remarkably resilient: a human lifetime of 70 years and a heart rate of 70 beats per minute equate to 2.6×10^9 heartbeats, each of which requires a reliable elastic response from the collagen-reinforced arterial walls (see *Collagen Derived Materials*[#]).

3. *Examples of Natural Fibrous Materials Based on β-sheets*

3.1 Silks

(*a*) *β-strands, β-sheets, and β-sheet crystals.* The *β*-strand is another simple, regular conformation that can be adopted by amino acid sequences in a protein. Individual protein chains with this conformation exhibit the following characteristics:

(i) The chains are almost fully extended, following a zigzag trajectory. If there is a twist, it is right-handed and gradual, ranging from an effectively infinite pitch in spider drag line silk to several hundred angstroms in many other protein crystals.

(ii) Stability of the conformation is maintained by hydrogen bonds with neighboring strands.

A planar arrangement of juxtaposed hydrogen-bonded *β*-strands looks like a pleated sheet, and is usually referred to as a *β*-sheet. Such sheets can stack into three-dimensional *β*-sheet crystals (Fig. 3), in which the separation of the *β*-sheets is dictated by the size of the side groups that must be accommodated within that space. The repeat distances in the covalent bond and hydrogen bond directions are independent of the amino acid sequence in the constituent protein.

(*b*) *Non-periodic lattice* (*NPL*) *crystals.* In some spider silks, the complicated amino acid sequence and the non-equilibrium conditions of fiber spinning make it unlikely that good matches will develop between adjacent sheets in the largest crystals. While the

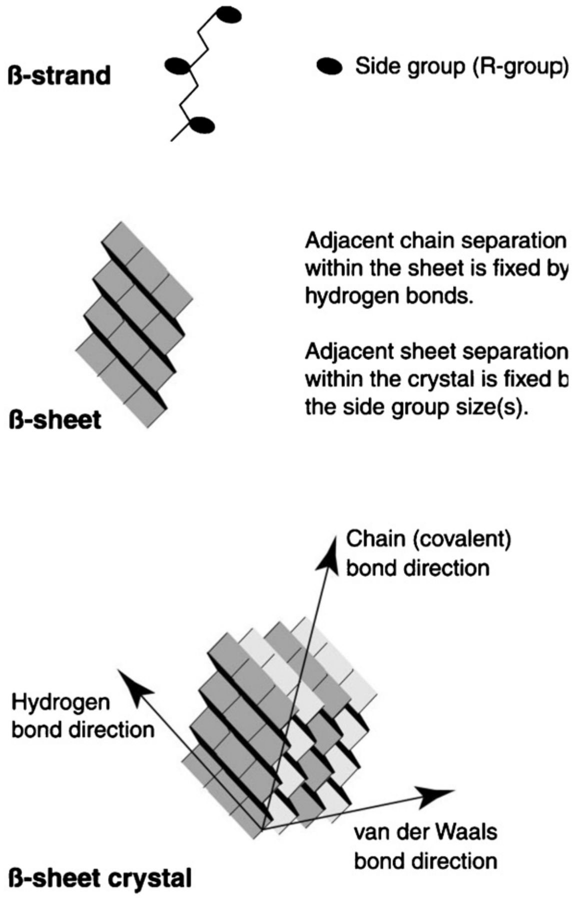

ß-strand
Side group (R-group)
Adjacent chain separation within the sheet is fixed by hydrogen bonds.
Adjacent sheet separation within the crystal is fixed b the side group size(s).
ß-sheet
Chain (covalent) bond direction
Hydrogen bond direction
van der Waals bond direction
ß-sheet crystal

Figure 3
Structure and principal features of a β-sheet crystal.

molecular backbones tend to pack on a three-dimensional lattice, this periodicity is frustrated by the complex sequence of amino acid side groups along each molecule, and the consequently necessary fluctuations of the inter-sheet spaces in which the side groups are accommodated.

The term ‘‘non-periodic lattice crystal’’ therefore succinctly describes the distinguishing features of this phase. Because NPL crystals involve statistical rather than perfect matches, they can reliably confer reinforcement to the solidified microstructure regardless of the exact processing conditions (spinning rate, temperature, humidity).

(*c*) *Properties.* Silks exhibit a wide range of optimized tensile properties, for example stiffness of the radial threads and frames of spider webs, compliance of the capture spiral in webs, and toughness of egg cocoons. Due to the extended conformation of their constituent chains, the β-sheet crystals confer axial strength and stiffness. Segments of the same polymer chains that build β-sheet crystals also make up the intercrystalline amorphous phase, so there is no discrete interface between these phases. The fiber can therefore extend significantly once the yield strength imparted by the crystalline phase is exceeded.

Hydrogen bonding and water are important to silk in several respects. Individual secreted protein molecules are folded into a globular (isotropic) conformation; this shape is stabilized by intramolecular hydrogen bonds which are balanced by the favorable interactions between water molecules and hydrophilic residues at the surface of the globules. At higher protein concentrations, the coils assemble into anisotropic structures. The driving force for this process has not been quantified, but may be due to the greater entropy achieved when bound water is released from the surface of coils and replaced with coil–coil interactions. The aqueous solution of linear aggregates forms a processable liquid crystalline phase from which spiders and silkworms spin water-insoluble fiber. As the solution is spun, the intramolecular hydrogen bonding pattern is disrupted, chains are extended, and hydrophobic residues are exposed to water; β-sheet crystals develop from the more regular amino acid sequences. Their hydrophobic side group content, dense molecular packing, and extensive hydrogen bonding between backbones makes the crystals insoluble in water.

Water also affects the macroscopic properties of silks. Spider drag line undergoes large, non-volumetric dimensional changes that depend on moisture content, allowing spiders to achieve the correct tension in the web framework as it is spun. The energy dissipation mechanisms available to capture thread are partly the result of its microstructure, and partly depend on the presence of water.

3.2 Viral Spike Protein

The structural characteristics of virus coats (capsids) are highly relevant to virus propagation, and are therefore the subject of intensive study. The coat must contain and protect the nucleic acid (genetic) contents, be resilient against impact, be capable of broaching the outer wall of a target cell, and provide a secure pathway for conducting nucleic acid into the target. Hollow spikes on the capsid fulfill the latter two roles, from which it has been deduced that they must have unusually high strength and stiffness in axial compression. Because compressive strength and stiffness have been a long-term but elusive goal of polymer science, the hierarchical structure of spike proteins deserves careful attention.

(*a*) *Cross-β-sheets.* Viral spike protein contains several repeats of relatively short β-strand-forming amino acid sequences. The chain folds back and forth to assemble a β-sheet, stabilized by intramolecular hydrogen bonds (Fig. 2). Because the long axis of the sheet is transverse rather than parallel to the molecular backbone, the structure in this case is known as a cross-β-sheet. Some silks, e.g., the egg stalks of green lacewing flies, are also reinforced by cross-β-sheet crystals.

(*b*) *Higher-order structure.* The structure varies significantly between different virus types. For both rotavirus and human adenovirus, there is evidence that three cross-β-sheets interact to form a hollow trimeric box beam that resists buckling in compression. Structural characterization is hampered by the small size of individual spikes (length $\leqslant 30$ nm and width $\leqslant 5$ nm). Hydrophobic bonding is presumed to be important in stabilizing their structure, since 70% of the amino acid residues are hydrophobic in representative cases.

Genetically engineered model polymers, based on multiple consecutive copies of the principal repeated sequence in native adenovirus spike protein, can self-assemble into a liquid crystalline phase in solution. However, trimeric box beams have not been found in the hierarchical microstructure of fibers spun from this phase. The spun fibers are formed under significantly off-equilibrium conditions, and should not be expected to have the same internal structure as the native spikes.

(*c*) *Properties.* In tensile tests conducted on dry fiber, the breaking strength is ~0.3 GPa, the stiffness ~4 GPa, and the elongation to failure ~ 30%. Given the evidence that the native spikes rely on hydrophobic bonding to maintain their structure, the properties of analog fibers as measured in air are likely to be inferior compared to results obtained in water—even if the detailed microstructure of the native material could be reproduced. In other words, the natural material is designed to work in an aqueous medium, and attempts to mimic its properties must take this reality into account.

4. Examples of Natural Fibrous Materials Based on Aggregation of Globular Molecules

4.1 Actin

(*a*) *Structure.* Actin is a principal constituent of muscle, and plays an important role in cell motility. Individual protein chains fold into a non-spherical globular conformation that can fit into a space approximately $5.5 \times 5.5 \times 3.5$ nm (Fig. 4). These globular molecules are conventionally called G-actin. They self-assemble into a right-handed, double-helical, elongated aggregate that is known as F-actin in recognition of its fibrous structure. The rate at which F-actin forms is limited by nucleation: non-covalent association between just two G-actin molecules is much weaker than the interactions binding G-actin in a larger aggregate, because there are more nearest-neighbor interactions per G-actin molecule in the latter case. Self-assembly of F-actin is driven by several interlinked factors:

(i) There are entropic considerations. G-actins that have become aggregated are able to bind significantly fewer water molecules compared to the same number of independent G-actins. Although G-actin aggregation is synonymous with an increase in G-actin order, the accompanying increase in the disorder of water is more than enough to compensate.

(ii) The F-actin rods can become locally oriented, forming liquid crystalline domains. Axial growth of the rods is favored within such domains, and longer rods are in turn more successful at stabilizing orientational order.

(iii) Actin-bundling proteins help to reduce the chemical activity of G-actin in solution, allowing a greater concentration to accumulate in a particular location and thus enhancing the chances of contact and aggregation.

(iv) Aggregation is promoted by an increased concentration of cationic solutes, because cation-phosphate complexing helps to reduce the activity of phosphate released by the aggregation reaction.

(v) The presence of globular proteins that are not attracted to actin can promote liquid crystalline phase separation, because the mixing of rods and flexible coils is entropically unfavorable.

(*b*) *Properties.* F-actin is rigid in tension, so that muscle and cytoskeletal fibers can slide past neighboring fibers without stretching. At the same time, actin fibers are flexible in torsion, allowing cells to change shape; this flexibility is provided by the hinged structure of the individual G-actin molecules (Fig. 4).

The ability of cohesive, durable fibrous structures to assemble via non-covalent aggregation of globular molecules has significant implications for commercial polymers. Traditional polymer science and engineering holds that fluid polymer can be processed into useful fibers only if there is a greater-than-critical level of chain entanglement. Watching a weightlifter in action, one cannot dispute the cohesivity of fibers made from non-covalently associated, non-interpenetrating G-actin coils. However, a route has yet to be found for making analogous fibers that function in a dry environment.

4.2 Tubulin

Microtubules in flagella and actin filaments in muscle exhibit similar features of hierarchical structure. The most striking difference is that the tubulin-based structures are hollow.

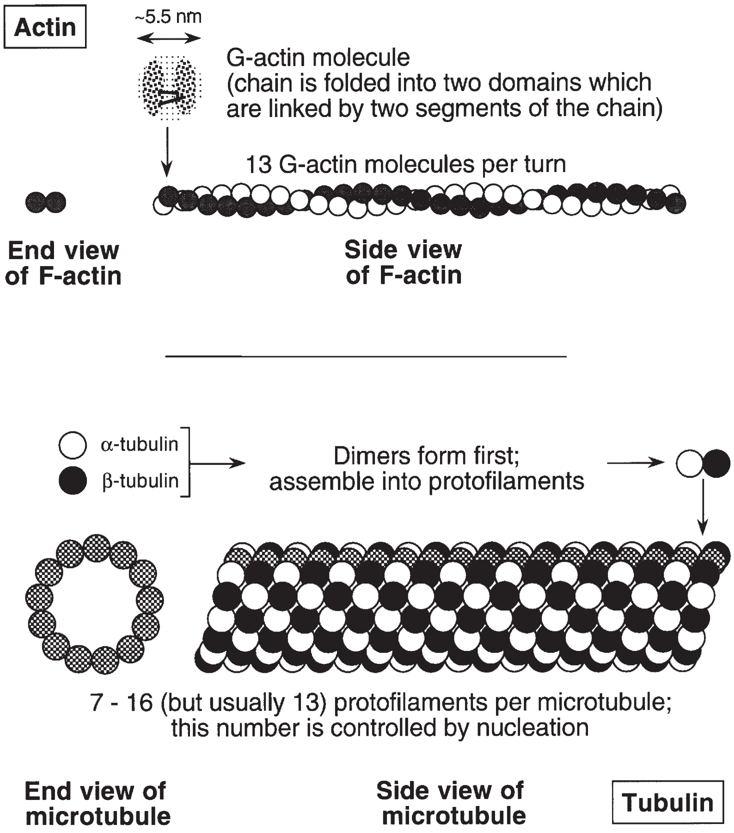

Actin
~5.5 nm
G-actin molecule
(chain is folded into two domains which are linked by two segments of the chain)
13 G-actin molecules per turn
End view of F-actin
Side view of F-actin
α-tubulin
β-tubulin
Dimers form first; assemble into protofilaments
7 - 16 (but usually 13) protofilaments per microtubule; this number is controlled by nucleation
End view of microtubule
Side view of microtubule
Tubulin

Figure 4
Molecular and supramolecular features of actin filaments and tubulin microtubules.

(*a*) *Dimers, protofilaments and microtubules.* Microtubules consist of two distinct types of globular protein, α-tubulin and β-tubulin (Fig. 4). The first stage of microtubule assembly involves α-tubulin/β-tubulin dimerization. A circular pattern of a third, related protein, γ-tubulin, occurs at the surface of MTOCs, providing a template to which the dimers are attracted and on which the microtubules are therefore nucleated.

A head-to-tail sequence of dimers forms a *protofilament*. The number of protofilaments in a microtubule can be as little as seven or as many as 16, depending on the *in vivo* or *in vitro* environment in which the microtubule is formed. There are usually 13 protofilaments in natural microtubules. Longitudinal bonds between tubulin molecules are stronger than latitudinal ones, leading to the view that microtubules grow by the simultaneous, longitudinal growth of all the constituent protofilaments. Whether or not the protofilaments are parallel to the microtubule axis depends on how many protofilaments there are. Parallelism is obtained with 13 protofilaments.

As was the case with actin, several factors combine to drive self-assembly. The number of protofilaments in the microtubule is dictated by the distribution of γ-tubulin in MTOCs. Each α/β-tubulin dimer exhibits a charge dipole, which helps it to assume the correct orientation to dock with a growing protofilament. There is evidence that liquid crystalline domains are formed in which axial growth and sustained alignment of rods are favored. The release of bound water from the aggregating tubulins can lead to a net increase of entropy in the solution.

(*b*) *Properties.* Like actin filaments, microtubules fulfill diverse roles. They are found in cytoskeletal fibers, flagella, cilia, and mitotic spindles, and are therefore integral to maintaining cell structure, to cell motility, and to cell reproduction. The fibers must be strongly cohesive to maintain rigidity under tension, despite the fact that they consist of non-covalently aggregated globular chains. The non-covalent bonds are sufficiently numerous, and collectively strong enough, for the globular actin molecules to maintain their shape under load and to remain connected to their neighbors in the aggregates. The tensile modulus of microtubules is approximately 1.2 GPa, similar to that of rigid plastics such as polymethylmethacrylate (PMMA) where significant molecular entanglements do occur. The flexural rigidity of microtubules is about 300 times greater than that of actin filaments. Microtubules are therefore rigid over cellular dimensions and can offer support to actin filaments in the cytoskeleton. Their rigidity is mainly due to their larger cross-section, achieved without unnecessary material expense because their cross-section is hollow.

5. The Future

Continuing studies of natural protein fibers will provide further insights into the principles of supramolecular materials self-assembly. The lessons learned will help to direct the development of new polymer fibers and other materials. Eventually, these lessons will lead to a practical understanding of how subtle chemistry (especially at the level of genetic coding) can be used to control the hierarchical structure and thence the properties of self-assembled materials and devices. *Adaptation* of Nature's lessons will be more fruitful than attempts to *copy* the structure or production route of any one natural material. Copying Nature is attended by the limitations of complex composition, slow growth, poor thermal stability, and sensitivity to the presence (or absence) of water.

See also: Collagen Derived Materials#

Bibliography

Ball P 1997 *Made to Measure: New Materials for the 21st Century*. Princeton University Press, Princeton, NJ

Elices M (ed.) 2000 *Structural Biological Materials*. Pergamon, Cambridge, UK

Hongu T, Phillips G O 1997 *New Fibers*. (2nd edn). Woodhead, Abington, UK

McGrath K, Kaplan D (eds.) 1997 *Protein-based Materials*. Birkhäuser, Boston, MA

National Research Council 1994 *Hierarchical Structures in Biology as a Guide for New Materials Technology*. National Academy Press, Washington, DC

Neville A C 1993 *Biology of Fibrous Composites*. Cambridge University Press, Cambridge, UK

Stryer L 1995 *Biochemistry*. Freeman, New York

Viney C 2000 From natural silks to new polymer fibers. *J. Textile Inst*. **91** (3), 2–23

Vogel S 1998 *Cats' Paws and Catapults: Mechanical Worlds of Nature and People*. Norton, New York

C. Viney
Heriot-Watt University, Edinburgh, UK

O

Optical Properties of Functional Hybrid Organic–Inorganic Nanocomposites

Functional hybrids are nanocomposite materials lying at the interface of organic and inorganic realms, whose high versatility offers a wide range of possibilities to elaborate tailor-made materials in terms of chemical and physical properties. Because they present several advantages for designing materials for optical applications (versatile and relatively facile chemistry, easy shaping and patterning, materials having good mechanical integrity and excellent optical quality), numerous silica- or/and siloxane- based hybrid organic–inorganic materials have been developed in the past few years. Figure 1 illustrates the interest of using multifunctional hybrid sol–gel-derived hybrid organic–inorganic coatings developed by several glass packaging industries in Japan and in Europe.

The most striking examples of functional hybrids exhibiting emission properties (solid-state dye lasers, rare-earth-doped hybrids, electroluminescent devices), absorption properties (photochromic), nonlinear optical (NLO) properties (second-order NLO, photochemical hole burning (PHB), photorefractivity), and sensing properties have been extensively reviewed (Avnir *et al.* 1994, Boilot *et al.* 1996, Dalton *et al.* 1995, Dunn and Zink 1991, Levy 1997, Sanchez and Lebeau 1996, 2001, Sanchez *et al.* 2003). These properties are summarized in this article.

Figure 1
Easy to recycle colored glass bottles coated by hybrid organic–inorganic materials (Minami 2000).

1. *Hybrids with Emission Properties*

There has long been an interest in fabricating solid-state gain media containing organic laser dyes, because it provides ease of use and replacement. Moreover, the encapsulation of the dye in solid matrices is favorable for decreasing health and environmental hazards.

Avnir *et al.* (1984) first synthesized luminescent dye-doped sol–gel matrices and demonstrated the possibility to use them as solid-state dye laser (see *Microchip Lasers*[#]). However, it is only a few years later that first laser emissions were reported in sol–gel materials (see *Sol–Gel Reactions: Theory*[#]). The main reasons of this delay are several drastic conditions required to elaborate efficient solid-state organic dye lasers. The laser material must exhibit an excellent transparency without scattering and a well-polished surface (roughness surface < few nanometers). Class II hybrid materials where organic and inorganic components are covalently linked were quickly found to be well appropriate for elaborating such solid-state dye lasers. Indeed, the organic component could improve matrix characteristics such as hydrophobicity to repel residual polar solvent and thus protect the organic dye from chemical degradation, density to decrease optical inhomogeneities, and mechanical resistance necessary to withstand surface polishing. Moreover, the covalent bonding of the dye allows a high dye loading to be achieved with minimized dye aggregation.

Performances of solid-state dye-laser materials are not only based on properties of the host matrix but also on the choice of the organic dye (see Fig. 2(a)). The dye must exhibit a good laser gain and must be chemically, photochemically, and thermally stable under the conditions of use of the laser.

The huge optimization potential of hybrid solid-state dye-laser materials is illustrated by the strong improvement of their performances over the past few years. Slope efficiency increased from a few percent to more than 80%. The maximum output energies reached simultaneously from a few μJ to a few mJ. At the same time, lifetimes progressed from about 10^4 pulses in the μJ range to more than 10^6 pulses in the mJ range.

Since the report of bright emission from thin layer molecular dye films, there is a growing interest in using organic electroluminescent devices in display applications because of their low operating voltage, large viewing angle, and high efficiency and brightness. Sol–gel processes provide the opportunity to design new hybrid light emitting diodes (LEDs) based on functionalized alkoxysilanes that are able to transport charges and emit light.

---

* Cross references marked by an asterisk are included in this volume.

[#] Cross references marked by a hash can be found by consulting the Encyclopedia of Materials: Science and Technology.

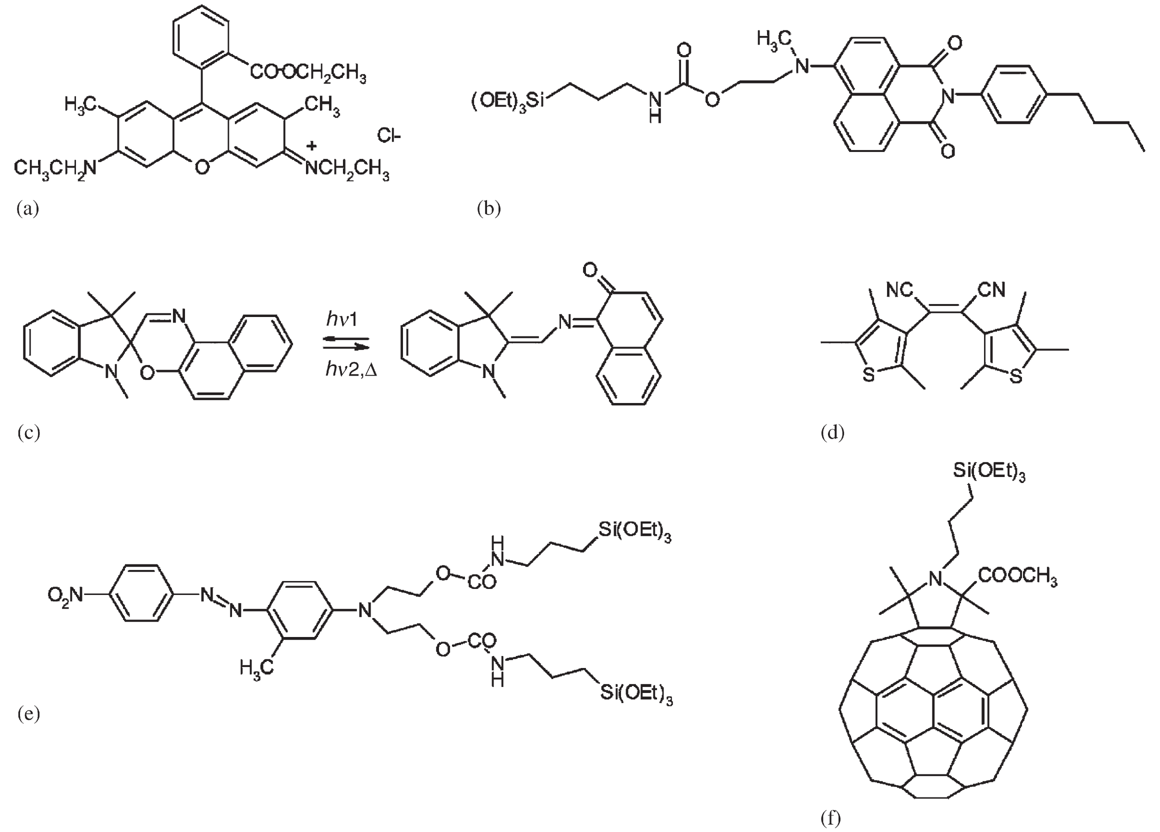


Figure 2
Molecular structures of laser-dye rhodamine 6G (a), silylated precursor with light emitting naphthalimide used to prepare electroluminescent devices emitting in the green (b), photochromic spirooxazine (SO) dye for fast optical switches (c), photochromic furylfulgide dye for optical data storage (d), silylated Red17 dye for nonlinear optics(e), and silylated fullerene C60 for optical limiters(f).

Hybrid organic–inorganic light emitting diodes (HLEDs) were elaborated for the first time by using new silylated precursors with hole transporting units prepared by modification of different active molecules (carbazole, oxadiazole, and tetraphenylphenylen-diamine derivatives) (see Fig. 2(b)).

All the first results on HLEDs are promising and further experiments are required to test these new materials as hole transporting layers in photorefractive and electroluminescent devices.

Due to their advantages, such as low-temperature processing and shaping, higher sample homogeneity and purity, and the opportunity to prepare noncrystalline solids, sol–gel matrices have been investigated for some years as potential matrices for lanthanide luminescence. Rare-earth ion (Nd^{3+}, Eu^{3+}, Er^{3+}, Tr^{3+}, Ce^{3+}, Eu^{2+},…) doped materials offer a wide range of optical applications such as tunable lasers, luminescent displays and amplifiers for optical communication.

In order to be able to achieve these objectives, the main strategies consist of the following: (i) Encapsulation of lanthanide ions within nanoparticles before embedding the doped nanoluminophore inside a hybrid matrix. (ii) One-pot sol–gel synthesis performed through a careful tuning of the sol–gel processing conditions (reactive hydrophobic precursors, hydrolysis ratio, catalysts, curing at moderate temperatures, etc.) to allow minimization of the hydroxyl content. (iii) Finally, the use of nonhydrolytic sol–gel processes permits it to generate hybrid matrices free from OH groups. Their main drawback is related to the difficulty to obtain from such processes transparent monolithic pieces or films (i.e., powders are usually obtained).

The sol–gel process provides also the unique opportunity to introduce in the same solid network both organic and inorganic chromophores. Moreover, the location of the organic dyes can *a priori* be tuned through dye–matrix interactions (covalent bonds, hydrogen bonds, ionic and van der Waals bonds). Only a few studies have reported on the simultaneous incorporation of rare-earth ion and organic dye into sol–gel-derived hybrid hosts. Energy transfer between Nd^{3+} and rhodamine 6G dyes incorporated inside hybrid siloxane oxide coatings was demonstrated, opening a multitude of possibilities to make new optical devices.

2. *Hybrid with Absorption Properties: Photochromic Hybrid Materials*

Photochromic materials can be used for the design of optical switches, optical data storage devices, energy-conserving coatings, eye-protection glasses, and privacy shields. Photochromic materials and systems have several important uses depending on the rates of the optical transformations. For example, very slow transformations are useful for optical data storage media, whereas fast transformations are required for optical switches.

Photochromic compounds such as diarylethene (DE), dithienylethene (DT), and furylfulgide (FF) derivatives are thermally stable and resistant to photochemical side reactions, were mainly used in optical information storage (see Fig. 2(d)).

Photochromic molecules such as spirooxazines (SO) and spiropyrans (SP) have been embedded into sol–gel glasses to produce optical switches (Fig. 2(c)). The rates of both light-driven structural changes and those of the back reactions are sensitive to the environment in the matrix. Pioneering work in this field by Levy and co-workers (summarized in Levy 1997) showed the important role played by the dye–matrix interactions on the photochromic response of spiropyrans. For hybrid organic–inorganic matrices containing different chemical environments, a competition between direct and reverse photochromisms can be observed. Several authors introduced the hydrophobic/hydrophilic balance (HHB) of the hybrid matrix as an important factor, which controls the competition between direct and reverse photochromism.

New nonsiloxane photochromic hybrid matrices made with spyrooxaxine or cationic spyropyran dyes embedded in nonsiloxane-based hosts have been reported very recently.

3. *Nonlinear Optics*

Quadratic nonlinear optics is a domain of "sol–gel optics" where the optical response and its temporal stability are strongly dependent on the nature of the bonding between the organic dye and the matrix. The random orientation of organic dyes in sol–gel matrices rules out second harmonic generation (SHG). Second-order nonlinearities can only be observed after poling in order to orient the organic dyes within the sol–gel matrix. Sol–gel processed hybrid materials, being usually amorphous, do not fulfill easily the phase-matching conditions required for frequency doubling. As a consequence, the most targeted applications are related to the electro-optic effect (electro-optic modulator, photorefractive materials, etc.).

The achievement of these materials is related to the quadratic susceptibility of the dye and the thermal stability of the polymeric hybrid matrix. Therefore, hybrid materials built from many matrix–NLO chromophore couples started to be optimized in terms of intrinsic NLO response, matrix rigidity, and dye–matrix interactions.

The best results are generally obtained when NLO dyes are grafted to the oxide matrix via several covalent bonds (hybrids of class II) (Dalton *et al.* 1995, Lebeau *et al.* 1997, Sanchez *et al.* 2003) (see Fig. 2(e)). Moreover, class II hybrid organic–inorganic materials allow a better control of the film thickness, the film hydrophobicity, the matrix rigidity, densification, and NLO chromophore concentrations. It was also observed that dye grafting leads to higher surface quality films and corona poling can be optimized to preserve this quality.

However, as far as NLO devices are concerned, measurements of electro-optical efficiency, wave guiding properties, and the evaluation of the optical losses must complete these promising results.

The large potential of these NLO hybrid materials in the field of photorefractive materials was demonstrated by Darracq *et al.* (1997). Photorefractive effects have been reported in hybrid sol–gel materials containing a second-order nonlinear chromophore (Disperse Red 1 (DR1) or 4-[N,N-bis(β-hydroxyethyl)amino]4′ nitrostilben (DHS) derivatives) and a charge transporting molecule (carbazole unit: CB) which was used to form a charge transfer complex with the photosensitizer (Trinitrofluorenone: TNF) to facilitate photocarrier generation at the visible wavelengths.

PHB has been attracting attention in recent years because of its scientific interest and the possibility of the application to high-density optical storage. Furthermore, PHB is a useful method to estimate the local structure around the organic dyes. Among several requirements for practical application, high hole efficiencies and high-temperature ($T > 77\,\mathrm{K}$) persistent spectral hole burning with long-term stability are the most important. PHB was first achieved in various organically modified silica matrices doped with different porphyrins or quinizarin derivatives.

Efficient hole burning was observed in hybrid matrices in which protoporphyrin IX was grafted by two chemical bonds to the inorganic network (Veret-Lemarinier *et al.* 1995). These hybrids of class II look promising for high-temperature persistent hole-burning materials.

The interest of hybrid matrices to preserve optical properties of a molecule was also shown with dyes known to exhibit strong reverse saturable absorption

(RSA). Efficient solid-state reverse saturable absorbers were made by entrapment of phtalocyanines and fullerene C60 in hybrid matrices. Improved LO performances were obtained by the use of grafting of C60 to the matrix backbone. Fullerene C60 derivative has been functionalized with Si tri-alkoxyde groups to increase its concentration in the final matrix without any clustering effects.

4. *Hybrid Optical Sensors*

Due to their attractive optical properties, a particularly intensive activity has recently characterized the area of reactive organically doped sol–gel materials for sensor applications. The advantage of sol–gel materials of being easy-to-process, combined with several other inherent advantages such as tunability of physical properties, high photochemical and thermal stability, chemical inertness, and negligible swelling both in aqueous and organic solvents, have resulted in an abundance of sol–gel optical sensors.

In some cases, the cracking of the sol–gel bulk material is a problem. In order to overcome this problem, organic–inorganic hybrid materials have been used as host matrices. Organically modified sol–gels not only allow an easy incorporation of probes but also an easy tailoring of properties such as mechanical properties, hydrophobicity, and refractive index. In some applications, silanization of the dye provides the advantage of a more homogeneous and well-defined distribution of the dye in the glass film. Moreover, the covalent coupling of the dye to the substrate allows it to prevent any dye leaching and thus to improve the sensor performance. An immobilization by chemical bonding is more appropriate to small indicators and sterically hindered analyte molecules. On the contrary, the entrapment by steric hindrance is more adapted for large indicator molecules and small fast diffusing analytes.

The sol–gel method offers the advantages of tailor-made porosity that is of paramount importance in sensing chemicals, because it may control the analyte diffusion and the leachability of the optical probe.

In the optical sensor area, the sol–gel method has also the advantage of easy shape modeling. This process may be used to fabricate thin glass films that have well-defined optical and physical characteristics, such as thickness, porosity, and refractive index. Moreover, the sol–gel method has the advantage to be compatible with fiber optic techniques that allow to design remote sensing systems, which are suitable for analyses in toxic or corrosive environments.

5. *Hierarchically Organized Hybrid Materials for Optical Applications*

Since the recent development of organic templating growth of materials, the elaboration of new types of hybrid photonic materials is emerging whose structure and function are organized hierarchically. Ordered periodic micro-, meso-, and macroporous materials allow the construction of composites with many guest types, e.g., organic molecules, inorganic ions, semiconductor clusters, or polymers. These guest/host materials combine high stability of the inorganic host system, new structure-forming mechanisms due to the confinement of guests in well-defined pores, and a modular composition. This could lead to new optical materials in, for example, switches, nonlinear optics, and lasers. Exploiting the periodic arrangement of the pore systems of mesoporous solids such as MCM-41 type could generate new optical devices.

Already in the early studies several proposals were made to apply such mesoporous systems as optical materials and highly specific chemical sensors. The self-assembled host, especially, allows appreciable control of mesoscale dimensions, symmetry, and orientational ordering of guest–host structures.

These composites can be structured hierarchically, leading to optical materials that have the desired optical properties due to the electronic states and molecular arrangements, and that have the desired light propagation of these complex structures of mesoporous materials.

The first reports of the optical properties of hybrid material whose structure and function are organized hierarchically showed that these solids are promising candidates for optical applications, including lasers, light filters, sensors, solar cells, pigments, optical data storage, photocatalysis, and frequency doubling devices (Stein *et al.* 2000, Wirnsberger and Stucky 2001).

6. *Conclusion*

Optically active hybrid solids are the fruit of multidisciplinary research in chemistry and physics. In particular, the progress in this field largely depends on the core competence of inorganic, polymer, and organic chemists, and illustrates the central role of chemistry in the development of advanced materials with unprecedented performances.

Photonic materials are perhaps the first systems to take advantage of the hybrid materials approach. Significant improvements in the properties of photonic materials may be attributed to three factors: (1) matrix improvements (mechanical integrity, transparency, thermal conductivity, reduced optical loss); (2) improvements in dye characteristics (chemical, thermal and photochemical stability, and high optical efficiencies); and (3) improved control of dye–matrix interactions.

There is no question that the development of hybrid photonic materials will continue in the years ahead because of both technological and scientific factors. The applications for dye-doped systems are, to some degree, limited by photochemical stability and improvements are clearly necessary if these materials are to emerge beyond the demonstration level.

The basic knowledge of organic–inorganic interaction and interfaces will play a major role for designing hierarchically ordered hybrid structures. These advanced materials which are now just appearing satisfy the requirements for a variety of applications and, in particular, optical communications and chemical sensing.

See also: Optical Materials from Organically Doped Sol–Gel Materials[#]; Nonlinear Optical Materials[#]; Glass: Sol–Gel Coatings[#]; Inorganic Materials, Sol–Gel Synthesis of[#]

Bibliography

Avnir D, Levy D, Reisfeld R 1984 The nature of the silica cage as reflected by spectral changes and enhanced photostability of trapped rhodamine 6G. *J. Phys. Chem.* **88**, 5956–9

Avnir D, Brauwn L, Lev O, Levy D, Ottolenghi M 1994 Organically doped sol–gel glasses: chemical sensors, enzymatic sensors, electrooptical materials, luminescent and photochromic materials. In: Klein L C (ed.) *Sol–Gel Optics: Processing and Applications*. Kluwer, Boston, pp. 539–82

Boilot J P, Chaput F, Gacoin T, Malier L, Canva M, Brun A, Levy Y, Galaup J P 1996 Solides organo-mineraux par sol–gel et applications à l'optique. *C.R. Acad. Sci.* **322**, 27–43

Dalton L R, Harper A W, Ghosn R, Steier W H, Ziari M, Fetterman H, Shi Y, Mustacich R V, Jen A K Y, Shea K J 1995 Synthesis and processing of improved organic second-order nonlinear optical materials for applications in photonics. *Chem. Mater.* **7**, 1060–81

Darracq B, Canva M, Chaput F, Boilot J-P, Riehl D, Levy Y, Brun A 1997 Stable photorefractive memory effect in sol–gel materials. *Appl. Phys. Lett.* **70**, 292–4

Dunn B, Zink J I 1991 Optical properties of sol–gel glasses doped with organic molecules. *J. Mater. Chem.* **1**, 903–13

Lebeau B, Sanchez C, Brasselet S, Zyss J 1997 Design, characterization and processing of hybrid organic–inorganic coatings with very high second-order optical nonlinearities. *Chem. Mater.* **9**, 1012–20

Levy D 1997 Photochromic sol–gel materials. *Chem. Mater.* **9**, 2666–70

Minami T 2000 Easily recyclable colored bottles by the sol–gel process. *Sol–Gel Sci. Technol.* **18**, 290–1

Sanchez C, Lebeau B 1996 Hybrid organic–inorganic materials with second-order optical nonlinearities synthesized via sol–gel chemistry. *Pure Appl. Opt.* **5**, 1–11

Sanchez C, Lebeau B 2001 Design and properties of hybrid organic–inorganic nanocomposites for photonics. *Mater. Res. Soc. Bull.* **26**, 377–87

Sanchez C, Lebeau B, Chaput F, Boilot J P 2003 Optical properties of functional hybrid organic–inorganic nanocomposites. In: Sanchez C, Gomez-Romero P (eds.) *Hybrid Functional Materials*. Wiley, Chap. 12

Stein A, Melde B J, Schroden R C 2000 Hybrid inorganic–organic mesoporous silicates—nanoscopic reactors coming of age. *Adv. Mater.* **12**, 1403–19

Veret-Lemarinier A V, Galaup J P, Ranger A, Chaput F, Boilot J P 1995 High temperature spectral hole burning on porphyrins grafted to sol–gel matrices. *J. Lumin.* **64**, 223–9

Wirnsberger G, Stucky G D 2001 Mesoporous and mesostructured materials for optical applications. *Chem. Mater.* **13**, 3140–50

C. Sanchez and B. Lebeau
Université Pierre et Marie Curie, Paris, France

Organic Aerogels

While inorganic aerogels have dominated the sol–gel literature since the 1950s, organic aerogels represent a relatively new class of nanoporous materials. During the 1990s organic aerogels having an ultrafine cell/pore size ($<100\,\mathrm{nm}$), high surface area ($>400\,\mathrm{m^2\,g^{-1}}$), and a solid matrix composed of interconnected colloidal-like particles, platelets, or fibers have been synthesized from resorcinol–formaldehyde, phenol–resorcinol–formaldehyde, phenol–furfural, melamine–formaldehyde, polyurethanes, polyisocyanurates, and polyureas. In certain cases these organic aerogels can be pyrolyzed in an inert atmosphere to give carbon aerogels in the form of monoliths, powders, microspheres, or thin-film composites. Ko (1998), Husing and Schubert (1998), and Fricke and Tillotson (1997) have all published review articles that contain sections on organic aerogels.

For information on related materials see *Inorganic and Inorganic–Organic Aerogels** and *Xerogels**.

1. Synthesis

The sol–gel polymerization of a multifunctional monomer (i.e., one that contains more than two reactive sites) in solution, leading to the formation of a cross-linked gel that can be transformed into an aerogel, extends beyond the traditional hydrolysis and condensation of silicon alkoxides. While many organic reactions can proceed through a sol–gel transition to form a cross-linked gel, only some can be converted into aerogels. For example, cross-linked, transparent acrylamide gels are commonly used for carrying out biological separations; however, these gels cannot be supercritically dried to form aerogels. The two major requirements for producing a gel that can be converted to an aerogel include:

(i) favorable polymer–solvent interactions during polymerization that prohibit macroscopic phase separation; and

(ii) a high degree of intra- and intermolecular cross-linking.

Pekala (1989) initially identified the aqueous polycondensation of resorcinol (1,3-dihydroxybenzene) with formaldehyde as a synthetic route to organic aerogels. In this system resorcinol is the multifuctional monomer which contains three reactive sites. The formaldehyde adds to the resorcinol ring to form hydroxymethyl groups that then cross-link to form a three-dimensional gel in which water resides in the pore structure. After a solvent exchange step, the gels are dried from supercritical carbon dioxide to form aerogels. If this same polymerization is attempted with phenol as the trifunctional monomer, macroscopic phase separation takes place and a precipitate is formed. Thus, the second hydroxyl group present on the resorcinol ring appears to be important in allowing gelation to take place prior to any macroscopic phase separation. Interestingly, mixtures of phenol and resorcinol can be polymerized with formaldehyde to form homogeneous gels, but only in cases where the mole ratio of phenol to resorcinol is less than approximately 3.

Although resorcinol–formaldehyde aerogels transmit light, these materials are dark red as a result of oxidation products formed during the polymerization. To circumvent this problem, Pekala *et al.* (1990) showed that melamine–formaldehyde oligomers could be cross-linked in water under acid conditions to form colorless, transparent gels. These gels were then supercritically dried from carbon dioxide to form aerogels. The optical clarity of melamine–formaldehyde aerogels is similar to that of silica aerogels.

While the above reactions demonstrated the feasibility of organic aerogels, they each required a solvent-exchange step from water to acetone to carbon dioxide. Pekala *et al.* (1995) demonstrated that a phenolic–furfural reaction conducted in *n*-propanol could also lead to organic aerogels. De Vos and Biesmans (1996) patented organic aerogels prepared from polyisocyanate-based reactions in organic solvents. In both cases, the cross-linked gels were placed directly in a pressure vessel and dried with supercritical carbon dioxide to form aerogels.

2. Structure–Property Relationships

Aerogels have a unique morphology in that *both* the covalently bonded clusters of the solid phase and the interconnected pores of the gas phase have nanometer-sized dimensions. This nanostructure can be engineered through control of the initial sol–gel polymerization conditions. For example, Pekala *et al.* (1992) have shown that the [resorcinol]/[catalyst] (R/C) mole ratio of the starting solution dictates the particle size, specific surface area, and pore size of resorcinol–formaldehyde (RF) aerogels. High R/C ratios (R/C ~ 200) give larger particles and lower surface areas than aerogels produced at low R/C ratios (R/C ~ 50). Saliger *et al.* (1997) showed that RF gels produced at very high R/C ratios (R/C ~ 1500) could even be dried subcritically using acetone as the pore fluid. The rather coarse gel structure is able to withstand the capillary pressures more efficiently than gels produced at low R/C ratios. Similar effects have been found in melamine–formaldehyde gels where the pH of the starting solution determines the surface area and transparency of the dried aerogel.

The structures and properties of organic aerogels are analogous to their inorganic counterparts (e.g., silica). These materials have continuous porosity, an ultrafine cell/pore size (< 100 nm), high surface area (400–1100 $m^2 g^{-1}$), and a solid matrix composed of spherical particles, platelets, or fibers with characteristic diameters of 10 nm. This nanostructure is responsible for the unique thermal, acoustic, optical, electrical, and mechanical properties of aerogels.

2.1 Thermal Properties

In monolithic organic aerogels the total thermal conductivity can be described by a simple addition of the gaseous thermal conductivity, λ_g, the solid thermal conductivity, λ_s, and the radiative thermal conductivity, λ_r:

$$\lambda_{total} = \lambda_g + \lambda_s + \lambda_r \tag{1}$$

In such systems coupling between the different heat transfer modes is negligible. Additionally, convection of the gas within the pores is completely suppressed.

The gaseous thermal conductivity in air-filled monolithic organic aerogels can be described by the Knudsen equation:

$$\lambda_g = \frac{\lambda_{g0}\Pi}{1 + \beta K_n} \tag{2}$$

where Π is the porosity and λ_{g0} is the thermal conductivity of free air. The parameter β takes into account the energy transfer between gas molecules and the limiting structure: for air, β is approximately 2. The Knudsen number

$$K_n = \frac{t_g}{\phi} \tag{3}$$

is a characteristic quantity for the gaseous thermal conductivity in a porous system, where t_g is the mean free path of the gas molecules and ϕ is the pore diameter.

Organic aerogels have pore sizes in the 1–100 nm range. The mean free path of the gas molecules in air at 1 bar is around 70 nm and thus is comparable to the average pore size in aerogels. Aerogels are such good thermal insulators because of this relative relationship. In the most simplistic model, there are only a few gas molecules trapped within an individual pore vs. many molecules which are transported from pore to pore as for example in Styrofoam. Thus, the gaseous conductivity, λ_g, is partially suppressed even at atmospheric pressure. Typical λ_g values range from $0.005\,\mathrm{W m^{-1} K^{-1}}$ to $0.010\,\mathrm{W m^{-1} K^{-1}}$, depending on density.

The above equations show that pore size is the important factor in reducing the gaseous thermal conductivity. Small average pore sizes increase the Knudsen number and decrease the gaseous thermal conductivity. For RF aerogels, smaller pores are achieved at a given density by reducing the R/C ratio.

Organic aerogels are optically thick for thermal radiation in layers of more than 1 cm, which means that the mean free path of infrared photons is small compared with specimen size. In this case radiative heat transport is a local phenomenon and the radiative conductivity can be described by:

$$\lambda_r = \frac{(16/3)n^2\sigma T^3}{e(T)\rho} \tag{4}$$

where σ is the Stefan-Boltzmann constant, n is the mean index of refraction of the specimen, T is the absolute temperature, $e(T)$ is the mass specific extinction coefficient and ρ is the density of the material. The extinction coefficient, $e(T)$, can be determined from infrared optical reflection and transmission experiments. Lu *et al.* (1995) found $e \approx 50\,\mathrm{m^2\,kg^{-1}}$ as being typical for RF aerogels. Because of the strong infrared absorption of all organic aerogel compositions, radiative heat transfer is more effectively reduced in these porous materials as compared with nonopaque silica aerogels.

The solid thermal conductivity of organic aerogels is drastically reduced with respect to the nonporous material. Furthermore, the solid conductivity depends upon the interconnectivity of the solid phase. Lu *et al.* (1995) found that the solid thermal conductivity scales with aerogel density as follows:

$$\lambda_s = C\rho^{\alpha} \tag{5}$$

The prefactor C depends upon the particle interconnectivity and $\alpha \approx 1.5$ holds for RF and carbon aerogels.

Figures 1 and 2 show the impact of the RF aerogel nanostructure upon gaseous and solid thermal conductivity, respectively. At a fixed aerogel density, the gaseous thermal conductivity is lower for samples prepared at R/C = 50 as compared to those produced at R/C = 200 or 300. This trend is a result of the smaller average pore size in organic aerogels synthesized at lower R/C ratios. In the case of the solid thermal conductivity, the opposite trend is observed. Those RF aerogels prepared at R/C = 50 exhibit a higher solid thermal conductivity at a given density as compared to those aerogels prepared at R/C = 200 or 300. This reflects the fibrous structure and better interconnectivity of RF aerogels synthesized at R/C = 50 as compared to the weak interconnection of larger particles with narrow necks for RF aerogels synthesized at R/C = 200 or 300.

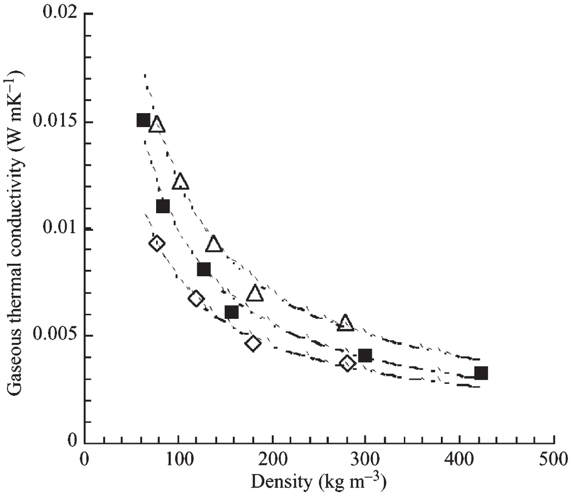


Figure 1
Dependence of gaseous thermal conductivity, λ_g, on the density of RF aerogels having different nanostructures: R/C = 50 (◇); R/C = 200 (■); R/C = 300 (△).

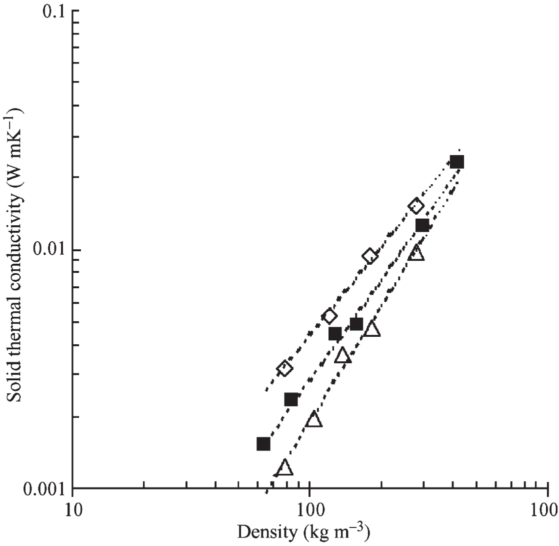


Figure 2
A log–log plot showing the dependence of solid thermal conductivity, λ_s, on the density of RF aerogels having different nanostructures: R/C = 50 (◇); R/C = 200 (■); R/C = 300 (△).

The total thermal conductivity of organic aerogels can be further suppressed if specimens are evacuated. Figure 3 shows the thermal conductivity of RF, polyisocyanurate, and carbon aerogels as a function of

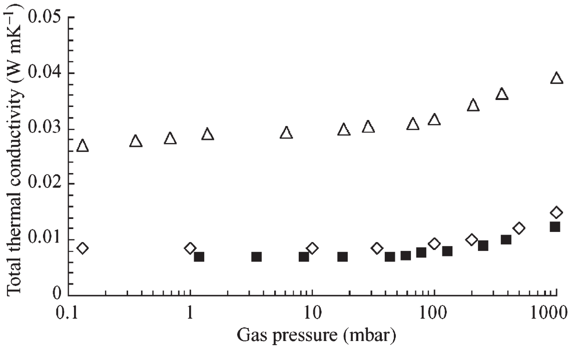


Figure 3
Total thermal conductivity of various types of organic aerogels as a function of gas pressure: C, $124\,\mathrm{kg\,m^{-3}}$ (△); PIR, $209\,\mathrm{kg\,m^{-3}}$ (◇); RF, $157\,\mathrm{kg\,m^{-3}}$ (■). The RF and carbon aerogels were synthesized at R/C = 200.

gas pressure. A drastic reduction in thermal conductivity is observed in all cases as the pressure is decreased below 100 m bar. This critical pressure defines where a loss in insulation would occur for monolithic aerogels used in vacuum insulation panels. Such panels are being considered for applications in household appliances such as refrigerators.

2.2 *Mechanical Properties*

Differences in the nanostructure of organic aerogels are also reflected in their mechanical properties. As a result of their high porosity, aerogels exhibit an elastic modulus many orders of magnitude smaller than their full-density analogs. Gross *et al.* (1997) and Pekala *et al.* (1990) have made extensive studies on the mechanical properties of RF and carbon aerogels. Gross *et al.* used acoustic techniques and a beam-bending method to evaluate the shear modulus, Young's modulus, and Poisson ratio of RF gels both before and after supercritical drying from carbon dioxide. They found that the elastic modulus typically increased by a factor of 1–5 after supercritical drying. This increase depended upon the initial gel concentration with the largest increases observed at the lowest gel concentrations. Thus, supercritical drying influences lower density gels more than high-density gels.

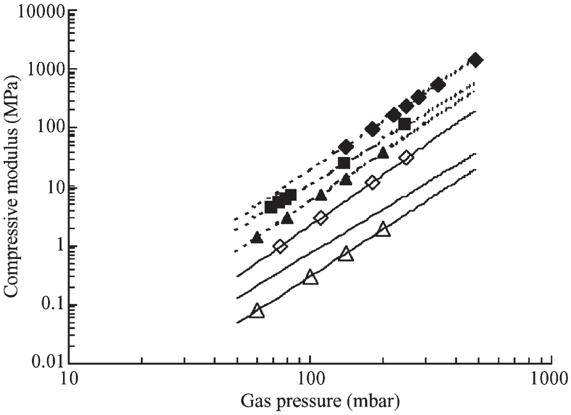


Figure 4
A log–log plot showing the dependence of compressive modulus on the density of RF (open symbols) and carbon (filled symbols) aerogels having different nanostructures: R/C = 50 (◇); R/C = 200 (■); R/C = 300 (△).

In analyzing the elastic modulus of RF aerogels over a density range of $40–1200\,\mathrm{kg\,m^{-3}}$, Gross *et al.* (1997) observed that a simple power-law fit could not be applied. In principle, the elastic modulus of porous materials scales with the second power of density, provided that the dominant deformation mechanism is bending of the beams or struts that make up the skeleton. When they limited the analysis to the density range $200–500\,\mathrm{kg\,m^{-3}}$, a scaling exponent of 3.5 was observed. This high exponent reflects the fact that organic aerogels are structurally inefficient. In other words, there are many "dead ends" within the aerogel network that contribute to the mass but not to the mechanical strength. Gross *et al.* (1997) observed that the fraction of elastically effective mass was as low as 17% for RF aerogels having a density of $\sim 100\,\mathrm{kg\,m^{-3}}$ and as high as 70% for RF aerogels having a density of $\sim 500\,\mathrm{kg\,m^{-3}}$.

Pekala *et al.* (1990) precision-machined RF and carbon aerogels into 1 cm × 1 cm × 1 cm cubes and tested their uniaxial stiffness (i.e., compressive modulus). These organic aerogels were synthesized at R/C ratios of 50–300, and all carbon aerogels were pyrolyzed at 1050 °C. Figure 4 shows a log–log plot of compressive modulus vs. density for these organic aerogel families. A scaling exponent of 2.7 ± 0.2 was observed for all formulations over the density range studied. Although the scaling exponents were identical, the modulus reflects the matrix material and R/C ratio used to produce the aerogels. At a fixed density, aerogels produced at low R/C ratios are stiffer than their high-R/C counterparts in either the RF or carbon families. These data reflect the smaller particle size and improved interconnectivity of aerogels produced at low R/C ratios.

A comparison of RF and carbon aerogels prepared at identical R/C ratios shows that the latter materials are an order of magnitude stiffer than their RF precursors. The improvement in stiffness increased with the R/C ratio. At equivalent densities, carbon aerogels were ~7 times stiffer than their RF counterparts at R/C = 50, ~16 times stiffer at R/C = 200, and ~23 times stiffer at R/C = 300. These data reflect a greater improvement in interparticle connectivity upon carbonization of RF aerogels prepared at

R/C = 300, which contain larger particles (~20 nm) with small interconnecting necks, as compared to RF aerogels prepared at R/C = 50 which already contain well-fused small particles or fibers (3–5 nm).

3. Applications

Aerogels have been proposed for many applications that take advantage of their unique structure and properties. Organic aerogels that exceed the performance of conventional foam technology in freezers or refrigerators would make it possible to offer more storage space or improved efficiency. Various companies (Aerojet, BASF, Hoechst, Armstrong World Industries, ICI) have demonstrated the feasibility of organic aerogel technology at the pilot scale; however, there are no organic aerogel-based vacuum insulation panels being used commercially. In fact, Ocellus Inc. (Alameda, CA) is the only commercial producer of organic aerogels.

Carbon aerogels represent the first electrically conductive aerogels. This property in combination with high surface area, controllable pore size, and high purity is leading to new electrochemical applications. Carbon aerogels are being investigated commercially as electrodes for electrochemical double layer capacitors (EDLCs) and electrosorptive processes such as capacitive deionization. Lipinski *et al.* (1998) have also examined the use of carbon aerogels as fuel cell electrodes.

Mayer *et al.* (1993) recognized that the high capacitance density (25–50 F cm^{-3}) of carbon aerogels would make them attractive for EDLC applications. In these devices the carbon aerogels serve as both the positive and negative electrodes, with a porous separator in between. The void space in both the separator and carbon aerogels is filled with an organic or aqueous electrolyte, and energy is stored via separation of charge across the polarized electrode/electrolyte interface. Such devices deliver higher power than conventional batteries, and they are finding applications in microelectronics and telecommunications. PowerStor Corp. (Dublin, CA) is manufacturing EDLCs based on carbon aerogel technology.

Capacitive deionization (CDI) can be viewed as an electrosorption process that takes place in a "flow-through" double layer capacitor. As contaminated water is passed through the CDI unit, charged impurities are transported to the electrode surface and held within the electric double layer to produce an effluent of purified water. Farmer *et al.* (1996) have studied the absorption profiles on many different ions captured on CDI units using carbon aerogel electrodes. This technology is being pursued commercially at Far West Corp. (Tucson, AZ).

Bibliography

De Vos R, Biesmans G 1996 Organic aerogels. *US Pat.* 5 484 818

Farmer J C, Fix D V, Mack G V, Pekala R W, Poco J F 1996 Capacitive deionization of NaCl and $NaNO_3$ solutions with carbon aerogel electrodes. *J. Electrochem. Soc.* **143** (1), 159–69

Fricke J, Tillotson T 1997 Aerogels: production, characterization, and applications. *Thin Solid Films* **297**, 212–23

Gross J, Scherer G W, Alviso C T, Pekala R W 1997 Elastic properties of crosslinked resorcinol–formaldehyde gels and aerogels. *J. Non-Cryst. Solids* **211**, 132–42

Husing N, Schubert U 1998 Aerogels—airy materials: chemistry, structure, and properties. *Angew. Chem., Int. Ed. Engl.* **37**, 23–45

Ko E 1998 Aerogels. *Kirk-Othmer Encyclopedia of Chemical Technology*. Wiley, New York

Lipinski M, Leuschner R, Gebhardt U, Petricevic R, Fricke J 1998 Kohlenstoff-Aerogels als dunne Gas diffusious Electrodes fur PEM Pressurstoffzeller, Werkstoffworche, Oct. 98. Munich, Germany.

Lu X, Caps R, Fricke J, Alviso C T, Pekala R W 1995 Correlation between structure and thermal conductivity of organic aerogels. *J. Non-Cryst. Solids* **188**, 226–34

Mayer S T, Pekala R W, Kaschmitter J L 1993 The aerocapacitor: an electrochemical double-layer energy-storage device. *J. Electrochem. Soc.* **140** (2), 446–51

Pekala R W 1989 Organic aerogels from the polycondensation of resorcinol with formaldehyde. *J. Mater. Sci.* **24**, 3221–7

Pekala R W, Alviso C T 1990 A new synthetic route to organic aerogels. *Mater. Res. Soc. Symp.* **180**, 791–6

Pekala R W, Alviso C T, LeMay J D 1990 Organic aerogels: microstructural dependence of mechanical properties in compression. *J. Non-Cryst. Solids* **125**, 67–75

Pekala R W, Alviso C T, LeMay J D 1992 Organic aerogels: a new type of ultrastructured polymer. In: Hench L L, West J K (eds.) *Chemical Processing of Advanced Materials*. Wiley, New York

Pekala R W, Alviso C T, Lu X, Gross J, Fricke J 1995 New organic aerogels based upon a phenolic–furfural reaction. *J. Non-Cryst. Solids* **188**, 34–40

Saliger R, Bock V, Petricevic R, Tillotson T, Geis S, Fricke J 1997 Carbon aerogels from dilute catalysis of resorcinol with formaldehyde. *J. Non-Cryst. Solids* **221**, 144–50

R. W. Pekala[a] and J. Fricke[b]
[a] *ENTEK International LLC, Lebanon, Oregon, USA*
[b] *Physikalisches Institut der Universistat Wurzburg, Germany*

Organic Analogues to Zeolites and Mesoporous Sieves

A major goal in the design of new heterogeneous catalysts has been to develop organic analogues to inorganic zeolites and mesoporous sieves. Zeolites are crystalline, open-framework materials with uniform channel diameters (3–10 Å) that have been used extensively in a number of industrially important processes, such as heterogeneous catalysis and size selective separations (Thomas *et al.* 1997) Mesoporous sieves are a relatively new class of surfactant-templated silicates pioneered by Mobil that contain adjustable pore sizes (2–10 nm) and architectures. They hold exceptional promise as new catalysts and separation media (Beck *et al.* 1992). The ultimate goal in developing organic analogues to these inorganic materials is to generate systems with the chemical tunability and processability of an organic material, and the high surface area, stability, and reactivity of a crystalline, supercage material. This article will provide an overview of methods for generating extended organic materials with open, nanometer-scale pore architectures mimicking those of the zeolites and mesoporous sieves.

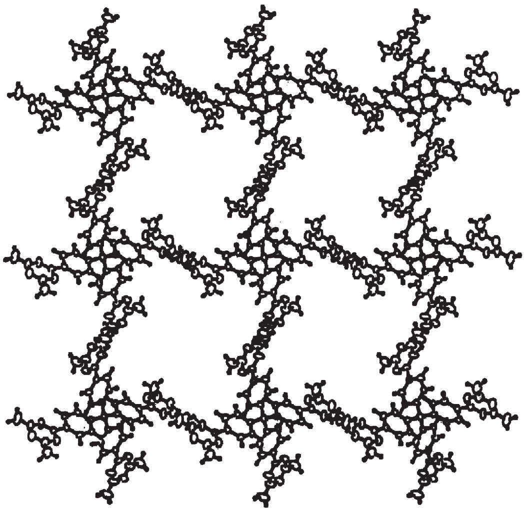
1. Formation of Porous, Noncovalent Organic Networks

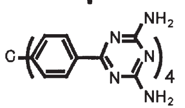
One of the most elegant approaches to the construction of nanoporous organic materials is the use of organic building blocks capable of forming specific noncovalent bonds. The resulting crystalline organic networks are capable of incorporating small molecules in the 5–12 Å periodic cavities formed upon their self-assembly from solution.

Several types of noncovalent bonding have been employed in forming such structures. One of the most prevalent is the hydrogen bond. Extended hydrogen-bond networks can be constructed by using molecules capable of forming multiple hydrogen bonds in specific directions. For example, Whitesides *et al.* (1995) formed a robust organic channeled structure by simply combining melamine and cyanuric acid to form the 1:1 adduct. Su *et al.* (1995) have used specially designed "molecular tecton" building blocks to form three-dimensional hydrogen-bond networks with adjustable porosity that can selectively include guest molecules.

A second type of noncovalent bonding motif that has been employed is ionic bonding in concert with hydrogen bonding. Swift *et al.* (1997) demonstrated that guanidinium organodisulfonates can be crystallized in the presence of organic guest molecules to produce bilayer structures in which the organodisulfonates serve as pillars that connect opposing two-dimensional hydrogen-bonded sheets of guanidinium and sulfate ions. The resulting one-dimensional pores in the material are occupied by the guest molecules. Pore dimensions can be tuned by the choice of organodisulfonate pillar and guest molecule.

These organic coordination networks provide excellent control over nanometer-scale pore structure; however, several major challenges still exist in their development as functional materials. The first is that many of these systems are not structurally robust toward collapse when the guest molecules are removed, due to the weak nature of the non-covalent interactions (Swift *et al.* 1997). However, structural integrity in the absence of guest molecules appears possible when the constituents are held in place by a large number of hydrogen bonds (Brunet *et al.* 1997) (Fig. 1). The second challenge is that the introduction of ancillary functional groups necessary for reactivity or selectivity in these systems is difficult to achieve without significant disruption of the desired lattice architecture (Swift *et al.* 1997). Consequently, only one example of catalytic activity in these organic coordination networks has been reported (Endo *et al.* 1997).





1

Figure 1
Crystal structure of the extended hydrogen-bond network of tecton **1** with formic acid. The formic acid molecules are omitted for clarity. (Reproduced by permission of the American Chemical Society from *J. Am. Chem. Soc.* **119**, 2737–8.)

2. Nanoporous Materials Via Phase-separated Block Copolymer Assemblies

A second approach for making nanoporous organic materials involves performing chemistry on phase-segregated block copolymers. Block copolymers are macromolecules produced by coupling two or more chemically and physically distinct polymer blocks. When the different blocks are sufficiently incompatible with one another, ordered phase-segregated architectures often spontaneously form. A variety of ordered nanometer-scale structures are possible from this self-organization process, including spheres, ordered cylinders, lamellae, and combinations thereof (Bates and Frederickson 1999).

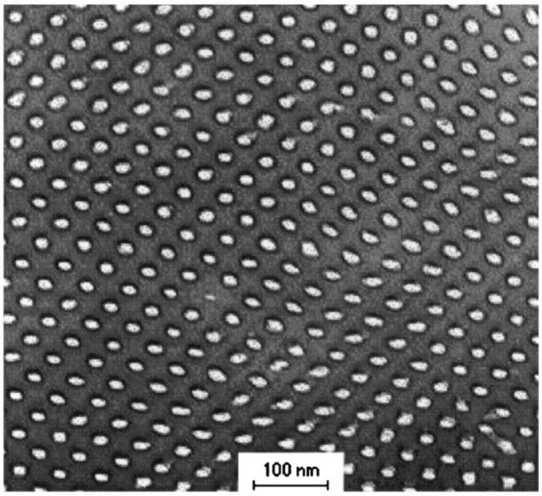


Figure 2
Transmission electron micrograph of a cross-linked diblock copolymer containing cylindrical cores of polyacrylic acid. (Reproduced by permission from Liu and Ding 1998.)

Lui and Ding (1998) have taken advantage of this phenomenon to generate nanoporous organic networks by using a diblock copolymer containing cross-linkable groups in one block and hydrolyzable groups in the other block. This material forms an ordered cylindrical phase in which the hydrolyzable block makes up the cores of the cylinders. Photopolymerization locks in the overall structure, and subsequent hydrolysis exposes hydrophilic carboxylic acid groups in the ca. 17 nm diameter cylinders (Fig. 2).

The most impressive features of these materials are their facile processability prior to cross-linking, their structural robustness due to cross-linking with covalent bonds as opposed to noncovalent interactions, and their ability to present certain functional groups in the pores. Catalysis in these nanoporous polymers has not yet been demonstrated but transport phenomena such as pH-gated water flow has.

3. **In Situ** *Polymerization of Surfactant-based Liquid Crystal Assemblies*

A third approach to the construction of nanostructured, porous organic assemblies is the cross-linking of self-assembled structures based on surfactant liquid crystals (LCs). Lyotropic or surfactant LCs are organic molecules that contain a hydrophilic headgroup and hydrophobic organic tail section. Above a certain concentration in water, these amphiphilic molecules spontaneously assemble into phase-segregated structures with the hydrophilic head-groups surrounding the water domains and the hydrophobic tails merging to form the hydrophilic domains. Depending on the shape of the surfactant LCs, the amount of water in the system, and the temperature, a range of nanostructured assemblies with periodic structures in the 1–10 nm range are possible, such as spherical micelles, hexagonal arrays of cylindrical micelles, lamellae, bicontinuous cubic phases, reverse micelles, and ordered arrays of rodlike reverse micelles (Tiddy 1980). By making polymerizable analogues of these surfactant LC phases, it is possible to lock in the structure of these assemblies by cross-linking with covalent bonds. Robust nanoporous polymer networks with water-filled pores are thus created in which hydrophilic functional groups are presented exclusively in the pore walls.

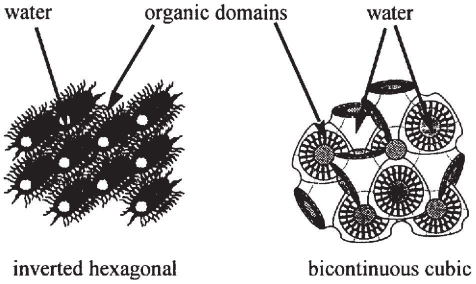


Figure 3
Schematic representations of the inverted hexagonal and *Pn3m* bicontinuous cubic phase of surfactant LCs. (Partially reproduced by permission of Elsevier Science. Copyright 1991.)

Early work by Menger and Tsuno (1990) in this area demonstrated the possibility of obtaining catalytically active, porous polystyrene networks by polymerizing a simple inverse micellar system consisting of cross-linker, water, and a polymerizable surfactant. In such an arrangement, the headgroups of the modified surfactant are localized at the inner surfaces of microscopic, spherical, "water pools." Anchoring of copper(II) ions to the headgroups in the cavities have enabled these porous networks to catalyze phosphate ester hydrolysis. Ruckenstein and Hong (1992) and Sundell *et al.* (1993) have used this "reactive microemulsion" approach to produce porous organic networks with anchored palladium centers to demonstrate catalytic vinylation and hydrogenation, respectively. Unfortunately, these materials do not possess the ordered pore system, high surface area, or uniform size pores characteristic of molecular sieves. These factors may limit their reactivity and selectivity.

Herz *et al.* (1963), Smith *et al.* (1997), and Srisiri *et al.* (1997) have independently succeeded in cross-linking the inverted hexagonal LC phase with retention of microstructure. In addition, Lee *et al.* (1995) have succeeded in cross-linking the *Pn3m* bicontinuous cubic LC phase with retention. The pore architectures in these systems (Fig. 3) are more similar to those of inorganic zeolites and mesoporous sieves, with nearly monodisperse pore diameters in the 1–7 nm range. Later studies have shown that different metal cations can be incorporated into the cross-linked inverted hexagonal phases to introduce new properties or alter pore dimensions (Deng *et al.* 1998). Preliminary studies have also demonstrated that cross-linked inverted hexagonal phases containing weakly basic sodium carboxylate headgroups in the nanochannels exhibit enhanced basicity and act as efficient heterogeneous base catalysts for organic reactions, with performances surpassing those of basic forms of zeolites and mesoporous sieves (Miller *et al.* 1999). The relevant features of this class of nanoporous organic materials are (i) structural tunability on the nanometer scale, (ii) processability prior to cross-linking, (iii) excellent structural integrity of the final product, and (iv) the ability to direct a variety of functional groups exclusively into the nanopores.

Bibliography

Bates F S, Frederickson, G H 1999 Block copolymers—designer soft materials. *Physics Today* (February), 32–8

Beck J S, Vartuli J C, Roth W J, Leonowicz M E, Kresge C T, Schmitt K D, Chu C T W, Olson D H, Sheppard E W, McCullen S B, Higgins J B, Schlenker J N 1992 A new family of mesoporous molecular sieves prepared with liquid crystal templates. *J. Am. Chem. Soc.* **114**, 10834–43

Brunet P, Simard M, Wuest J D 1997 Molecular tectonics. Porous hydrogen-bonded networks with unprecedented structural integrity. *J. Am. Chem. Soc.* **119**, 2737–8

Deng H, Gin D L, Smith R C 1998 Polymerizable lyotropic liquid crystals containing transition-metal and lanthanide ions: architectural control and introduction of new properties into nanostructured properties. *J. Am. Chem. Soc.* **120**, 3522–3

Endo K, Koike T, Sawaki T, Hayashida O, Masuda H, Aoyama Y 1997 Catalysis by organic solids. Stereoselective Diels–Alder reactions promoted by microporous molecular crystals having an extensive hydrogen-bonded network. *J. Am. Chem. Soc.* **119**, 4117–22

Herz J, Reiss-Husson F, Rempp P, Luzzatti V 1963 Quelques exemples de polymérisation en phase mésomorphe. *J. Polym. Sci. C.* **4**, 1275–90

Lee Y-S, Yang J-Z, Sisson T M, Frankel D A, Gleeson J T, Aksay E, Keller S L, Gruner S M, O'Brien D F 1995 Polymerization of nonlamellar lipid assemblies. *J. Am. Chem. Soc.* **117**, 5573–8

Liu G, Ding D 1998 Diblock thin films with densely hexagonally packed nanochannels. *Adv. Mater.* **10**, 69

Menger F M, Tsuno T 1990 Cross-linked polystyrene incorporating water pools. *J. Am. Chem. Soc.* **112**, 1263–8

Miller S A, Kim E, Gray D H, Gin D L 1999 Heterogeneous catalysis with cross-linked lyotropic liquid crystal assemblies: organic anologues to zeolites and mesoporous sieves. *Angew. Chem. Int. Ed.* **38**, 3021–6

Ruckenstein E, Hong L 1992 Binding catalytic sites to the surface of porous polymers and some catalytic applications. *Chem. Mater.* **4**, 122–7

Smith R C, Fischer W M, Gin D L 1997 Ordered poly(*p*-phenylenevinylene) matrix nanocomposites via lyotropic liquid-crystalline monomers. *J. Am. Chem. Soc.* **119**, 4092–3

Srisiri W, Sisson T M, O'Brien D F, McGrath K M, Han Y Q, Gruner S M 1997 Polymerization of the inverted hexagonal phase. *J. Am. Chem. Soc.* **119**, 4866–73

Su D, Wang X, Simard M, Wuest J D 1995 Molecular tectonics. *Supramolecular Chem.* **6**, 171–8

Sundell M J, Pajunen E O, Hormi O E, Näsman J H 1993 Synthesis and use as a catalyst support of porous polystyrene with bis(phosphonic acid)-functionalized surfaces. *Chem. Mater.* **5**, 372–6

Swift J A, Russell V A, Ward, M D 1997 *Adv. Mater.* **9**, 1183

Thomas J M, Bell R G, Catlow C R A 1997 Zeolites and related sieves. Organoporous hosts with adjustable molecular environments. In: Vol. 1, Ertl G, Knozinger H, Weitkamp J (eds.). *Handbook of Heterogenous Catalysis.* Wiley-VCH, Weinheim, Germany pp. 286–387

Tiddy G D T 1980 Surfactant-water liquid crystal phases. *Phys. Rep.* **57**, 2

Whitesides G M, Simanek E E, Mathias J P, Seto C T, Chin D N, Mammen M, Gordon D M 1995 Noncovalent synthesis: using physical–organic chemistry to make aggregates. *Acc. Chem. Res.* **28**, 37–44

D. L. Gin and S. A. Miller

University of California, Berkeley, California, USA

Organic–Inorganic Composite Crystals

Organic–inorganic hybrids offer the possibility of integrating organic and inorganic characteristics within a single "molecular composite" (Day and Ledsham 1982). Organic materials are generally lightweight, mechanically flexible, chemically versatile (e.g., molecules of different length, width, conformation), and can provide useful optical and electrical properties. Inorganic materials offer mechanical robustness, good thermal stability, high electrical and thermal conductivity, and interesting magnetic or ferroelectric transitions. Nature provides many examples of organic–inorganic composites that exhibit superior properties compared with those of the constituent organic or inorganic components (e.g., seashell, bone, tooth enamel).

A number of useful noncrystalline composites have also been constructed using techniques such as molecular beam epitaxy (MBE) (Takada 1995) and sol–gel processing (see, for example, Sanchez *et al.* 1999). This article will focus on those organic–inorganic hybrids that self-assemble from solution or vapor phase into a *crystalline* form, with mixing and interaction of the organic and inorganic components on a molecular level. The ability to grow the materials as single crystals enables structural studies and therefore the consideration of structure–property correlation. In addition, the structural self-assembly enables simple thin-film processing techniques and creates the potential for low-cost technological applications.

1. Crystal Structures

One scheme for classifying crystalline organic–inorganic composites involves sorting the hybrids based on the interaction between the inorganic framework and the organic moieties. The weakest interaction occurs in intercalated compounds, which are formed by reversibly incorporating organic molecules (guest species) into the interlayer or interchain gallery of a pre-existing host structure. Intercalated compounds have largely been studied in layered silicates and transition metal oxides, chalcogenides, and phosphates (Whittingham and Jacobson 1982). The structure of the inorganic framework is little affected by the guest molecule, highlighting the small degree of interaction between the two components. The next group comprises composites with stronger hydrogen and/or ionic bonding between the two components and includes organic–inorganic perovskites and related metal halide-based hybrids (Mitzi 1999). In contrast to the previous group, these compounds have specific stoichiometries, with the organic ions occupying well-defined sites in a crystal lattice. Finally, there are hybrids with strong covalent bonding between the two components. While this category includes all organometallic compounds (which generally consist of isolated organometallic clusters interacting via weaker molecular forces), several particularly interesting examples consist of extended inorganic frameworks with covalently bonded organic sidegroups (e.g., $C_nH_{2n+1}BiI_2$, CH_3ReO_3, $Zr(HPO_4)(C_6H_5OPO_3)\cdot 2H_2O$).

While each class of hybrids offers unique opportunities, the ionically bonded metal halide-based systems are particularly notable because of the high degree of ambient-temperature solubility of the inorganic framework in common solvents and the tendency of these compounds to crystallize or self-assemble from solution or the vapor phase. The organic–inorganic perovskites are one of the most extensively studied families of metal halide-based hybrids.

1.1 Perovskites

The basic motif of the organic–inorganic perovskite family is the AMX_3 structure shown in Fig. 1(a). This framework consists of a three-dimensional (3D) network of corner-sharing MX_6 octahedra, where the M atom is typically a divalent metal and X is a halide (Cl^-, Br^-, I^-). The A cations fill the large 12-fold coordinated sites between the octahedra and balance the charge of the MX_3^- network. One example of a 3D organic–inorganic perovskite is $CH_3NH_3SI_3$, where $CH_3NH_3^+$ comprises the simple organic structural component of the hybrid and SnI_3^- forms the 3D extended inorganic network.

Layered perovskites are conceptually formed by taking an *n*-layer-thick cut from the 3D perovskite structure along one crystallographic direction (e.g., $\langle 100 \rangle$ or $\langle 110 \rangle$) and alternating these layers with an organic modulation layer (Fig. 1(b)). Changing the thickness of the perovskite sheets alters the effective dimensionality of the inorganic framework, with a more 3D structure for large *n* and a more two-dimensional (2D) structure for small *n* (Fig. 2). In the $\langle 110 \rangle$-oriented family, the $n=1$ structure reduces to layers of one-dimensional (1D) chains. This enables a progression from 1D ($n=1$) to 2D ($n \geqslant 2$) to 3D ($n=\infty$) inorganic networks, all within a single structural family. With an appropriate choice of organic and inorganic components, these compounds resemble multilayer quantum well structures (typically produced using MBE or organometallic chemical vapor deposition (OMCVD)), with semiconducting inorganic sheets or chains (wells) alternating with much wider bandgap organic layers (barriers). However, the perovskites self-assemble into single-crystal structures and the effect of interfacial roughness between the well and barrier layers is therefore avoided. Modifying the length and/or the degree of conjugation within the organic layer molecules can also influence the electronic or magnetic interaction between inorganic layers.

Many studies have focused on the single layer ($n=1$) $\langle 100 \rangle$-oriented perovskites of the form $(RNH_3)_2MX_4$ or $(NH_3RNH_3)MX_4$, with M = Cu, Cr, Mn, Fe, Cd, Ge, Sn, Pb, or Eu; X = Cl, Br, I (Mitzi 1999). The monoammonium cations hydrogen bond to only one inorganic layer, leading to a bilayer of organic cations between the perovskite sheets and a van der Waals gap between the organic layers. For diammonium cations, the organic cations span the entire distance between and bond to two adjacent inorganic sheets. Generally, simple R groups have been incorporated, including alkane and single/double ring aromatic components. Functional dye molecules have also been included (e.g., Fig. 3(a)). In addition, unsaturated monomers, including diene and diyne derivatives, have been substituted and can, given the correct pairing of inorganic framework and organic monomer, undergo topochemical polymerization in the solid state upon ultraviolet (UV) light or γ-irradiation (e.g., Fig. 3(b)). The inorganic framework in these systems acts as a template to control the orientation and conformation of the organic molecules, thereby influencing the degree to which the organic layer polymerizes upon irradiation.

1.2 Other Systems

While the perovskites are among the most extensively studied organic–inorganic hybrid crystal structures, with their characteristic framework of corner-sharing metal halide octahedra, alternate inorganic frameworks can be built up from edge-sharing or face-sharing octahedra (e.g., Corradi *et al.* 1997). In some

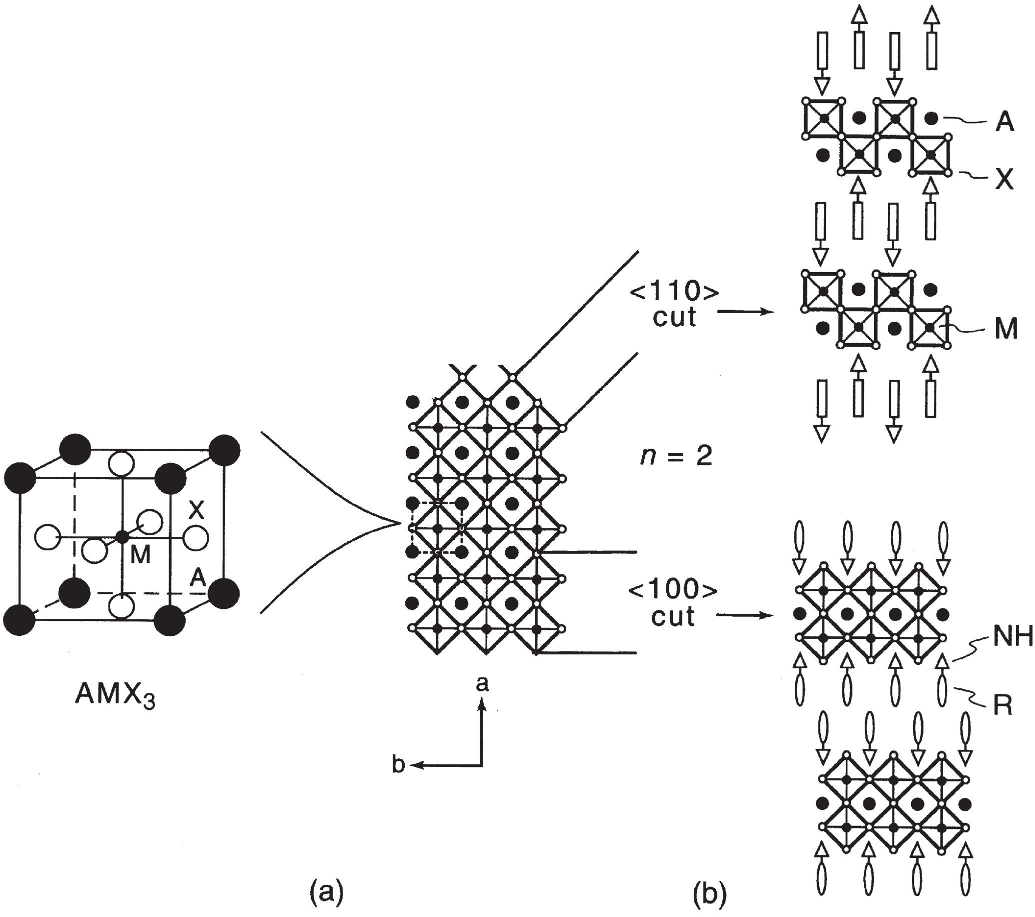


Figure 1
(a) The basic AMX_3 perovskite structure and a schematic representation of how the structure extends in three dimensions. The dashed square corresponds to the unit cell on the left. (b) Two families of layered perovskites, which are formed by taking *n*-layer-thick slices along the $\langle 100 \rangle$ and $\langle 110 \rangle$ directions of the AMX_3 structure, respectively, and stacking these layers alternately with an organic bilayer. In each case, only the $n = 2$ structure is shown. Different organic cations are used to stabilize the $\langle 100 \rangle$- vs. the $\langle 110 \rangle$-oriented structures.

cases, the repeating motif is not an octahedron, but rather one of many other possible geometrical forms, including tetrahedra, trigonal bipyramids, or dumbbells. One interesting example (Fig. 4) involves 2-amino-5-nitropyridinium cations ionically bonded within a layered inorganic framework of hydrogen-bonded $H_2AsO_4^-$ tetrahedra (Pecaut *et al.* 1993). The inorganic framework acts to screen the dipolar interactions between organic cations in adjacent layers and enables the formation of a polar structure that might be useful for quadratic nonlinear optical applications. As for the metal halide perovskites, $C_5H_6N_3O_2 \cdot H_2AsO_4$ crystals and films can readily be formed.

2. *Processing*

The interactions in the hybrid perovskites that promote self-assembly include covalent/ionic interactions, which favor the formation of sheets of corner-sharing metal halide octahedra, ionic/hydrogen bonding interactions between the organic cation and the inorganic sheet halogens, and the tendency of the organic cations to align "tail-to-tail." Other composite crystals exhibit a similar range of interactions. In general, self-assembling materials can be made by simple techniques, such as the Langmuir–Blodgett process (Petty 1996), since the various forces acting between the components favor the final atomic and

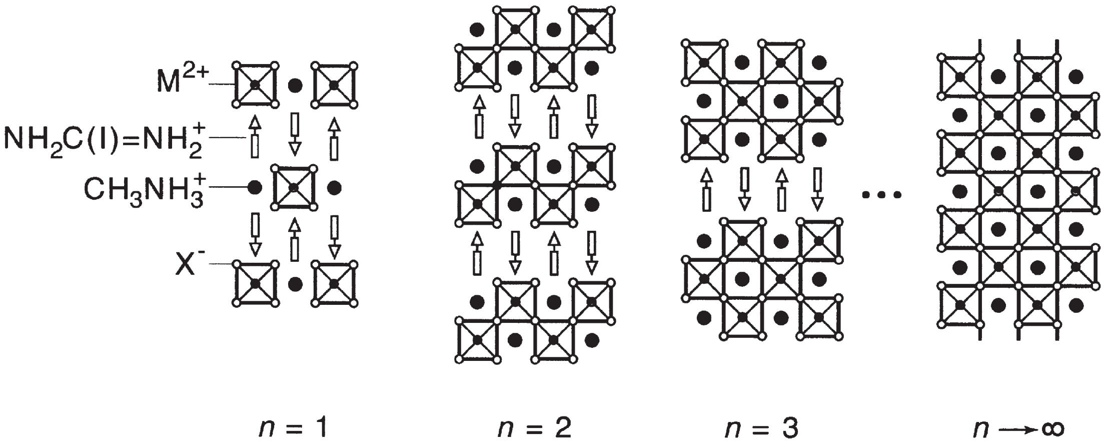


Figure 2
Schematic representation of a ⟨110⟩-oriented organic–inorganic perovskite family. In this case the ⟨110⟩-oriented structure is stabilized with the iodoformamidinium cation, leading to the family $[NH_2C(I){=}NH_2]_2(CH_3NH_3)_nM_nX_{3n+2}$, where M is a divalent metal and X is a halide (Cl^-, Br^-, I^-).

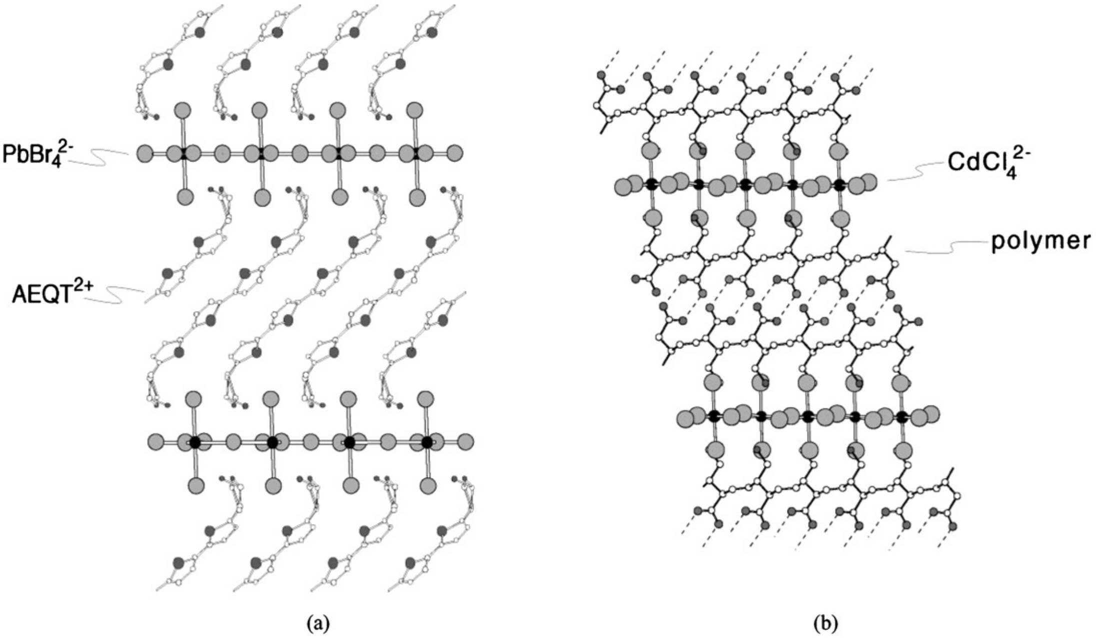


Figure 3
(a) Crystal structure of the dye-containing $n = 1$ ⟨100⟩-oriented perovskite (AEQT)$PbBr_4$, where AEQT is the doubly protonated form of 5,5‴-bis(aminoethyl)-2,2′:5′,2″:5″,2‴-quaterthiophene (Mitzi *et al.* 1999a). Note that in this structure, the organic cation is a diammonium cation and therefore bonds to two adjacent inorganic sheets. (b) Crystal structure of the polymerized product formed by subjecting $(HO_2CCH{=}CHCH{=}CHCH_2NH_3)_2CdCl_4$ to UV or γ irradiation. Polymerization yields ordered 1,4-disubstituted *trans*-polybutadiene $[–CH(CO_2H)CH{=}CHCH(CH_2NH_2)–]^{\infty}$. Hydrogen bonding between CO_2H groups on adjacent layers is shown by dashed lines (Tieke and Chapuis 1986).

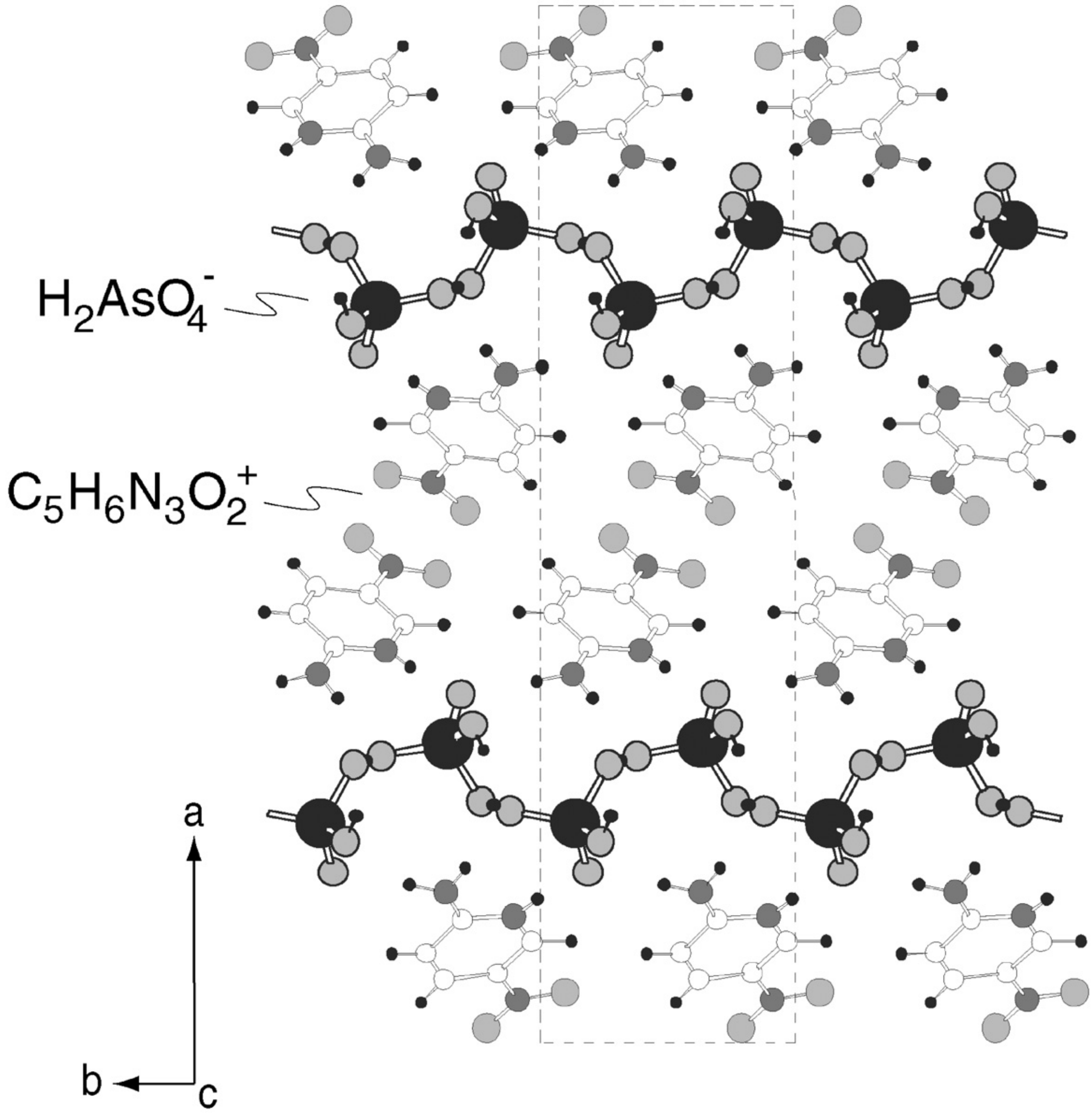


Figure 4
Crystal structure of $C_5H_6N_3O_2 \cdot H_2AsO_4$, viewed down the crystallographic c-axis. The dashed lines represent the orthorhombic (*Pna2₁*) unit cell outline (Pecaut *et al.* 1993).

molecular configuration. A number of convenient techniques can therefore be used to grow crystals and films of organic–inorganic composites—especially those based on a metal halide framework.

2.1 Crystal Growth

The relatively low decomposition temperature for most organic cations suggests that low-temperature solution crystal growth techniques are favored over higher-temperature techniques (e.g., growth from a melt). Solution crystal growth involves finding a suitable solvent for the component metal halide and organic salts. Several potential solvents include the relatively polar solvents, water, hydrohalic (X = Cl, Br, I) acids, various alcohols, acetonitrile, acetone, glacial acetic acid, ethylene glycol, and *N*-methyl-2-pyrrolidone. Generally, stoichiometric quantities of the organic salt and the metal halide are dissolved in the solvent and crystals are grown either through unseeded isothermal evaporation or through unseeded or seeded cooling of saturated solutions (Arend *et al.* 1978). While for many metal halide frameworks and simple organic cations a range of solvents are appropriate, for large, rigid, and relatively complex organic cations the competing solubility requirements of the organic and inorganic components of the structure limit the choice of solvent. For certain hybrids, layered solution growth and gel growth techniques have also been employed. The various techniques used for growing organic–inorganic perovskite crystals have been discussed (Mitzi 1999). Other metal halide-based crystalline composites can be similarly prepared.

2.2 *Thin-film Deposition*

While single crystals are generally the most useful medium for examining structural and physical properties of the hybrids, many applications and some measurements (e.g., optical properties) require the ability to deposit thin films. Crystalline organic–inorganic hybrids based on metal halide inorganic frameworks are deposited using a number of simple solution and evaporation techniques.

(*a*) *Solution techniques*. Spin coating enables thin-film deposition of a number of soluble organic–inorganic crystalline composites on a variety of substrates including glass, quartz, sapphire, and silicon. The process involves finding a suitable solvent for the hybrid, preparing a solution with the desired concentration, and applying drops of the solution to a spinning substrate. As the solution spreads on the substrate, the solution dries and leaves a deposit (generally crystalline) of the hybrid. The relevant parameters for the deposition include the choice of substrate, the solvent, the concentration of the hybrid in the solvent, the substrate temperature, and the spin speed. Pretreating the substrate surface with an appropriate adhesion promoter may improve the wetting properties of the solution. Postdeposition, low-temperature annealing ($T < 250\,^{\circ}\mathrm{C}$) of the hybrid films is also sometimes employed to improve crystallinity and phase purity. In addition to spin coating, other solution techniques for depositing the organic–inorganic hybrids include inkjet printing, stamping, and spray coating. All of these techniques require a suitable solvent for the organic–inorganic hybrid.

(*b*) *Evaporation techniques*. In some cases, a solvent cannot be identified for solution deposition of hybrid films. This could be due either to a lack of solubility of the hybrid or to surface wetting considerations for the given substrate. In addition, some applications or studies require deposition under vacuum conditions (e.g., thermal evaporation). The problem with evaporating the hybrid materials is that the organic component of the structure will generally dissociate or decompose from the material at substantially lower temperatures, or in a shorter time interval, relative to the inorganic component. Therefore, if a source containing the hybrid is gradually heated up in a vacuum chamber, the deposited films will not, in general, consist of the same hybrid as the initial source. Two approaches have been employed to overcome this thermal or temporal incompatibility between the organic and inorganic components during evaporation.

In two-source thermal evaporation, the organic salt and the metal halide components of the hybrid are contained in separate evaporation sources and each is heated appropriately to enable a nominally controlled evolution. Assuming that the evaporation rate for the two components can be controlled and calibrated, a stoichiometric and often crystalline film of the organic–inorganic composite can be prepared. Experiments have demonstrated (Era *et al.* 1997) the technique for the organic–inorganic perovskite $(C_6H_5C_2H_4NH_3)_2PbI_4$.

While the two-source technique is suitable for some hybrids and is entirely vacuum-compatible, it has the disadvantage of requiring control over the deposition rate of the organic salt. Often, while the metal halide evaporates in a well-defined fashion, the organic salt deposition is difficult to control. For each new organic–inorganic hybrid, new heater parameters are required to balance the rates of the two components. In addition, given the relatively long times required for evaporation, high vacuum conditions are needed during deposition.

A single-source thermal ablation (SSTA) technique has also been described (Mitzi *et al.* 1999b) which employs a single evaporation source and very rapid heating. A deposit of the organic–inorganic hybrid is placed on a thin tantalum heater within a vacuum chamber. When a large current is passed through the heater, the hybrid ablates over a very short interval ($<1\,\mathrm{s}$) and deposits on the substrates positioned above the heater. Since the ablation process is essentially instantaneous, both the organic and inorganic components evaporate from the heater at essentially the same time and, in general, without substantial decomposition of the organic component. Surprisingly, in many cases (especially with relatively simple organic cations) the as-deposited films are single phase and crystalline, indicating that the organic–inorganic hybrids can reassemble on the substrate at room temperature. For more complex organic cations, a short ($<20\,\mathrm{min}$), low-temperature ($<200\,^{\circ}\mathrm{C}$) anneal is sometimes required to improve film crystallinity.

(*c*) *Dip processing*. Some organic components will not withstand heating—even the short burst of energy leading to ablation during SSTA. In these cases, or if a suitable solvent for the organic–inorganic hybrid cannot be found, other techniques are required for thin-film deposition. A dip-processing technique (Liang *et al.* 1998) has been described in which a predeposited metal halide film is dipped into a solution containing the organic cation. The solvent for the dipping solution is chosen so that the organic salt is soluble, but the starting metal halide and the final organic–inorganic hybrid are not substantially soluble. The organic cations in solution react with the metal halide on the substrate and form a crystalline film of the desired hybrid. For the family $(RNH_3)_2(CH_3NH_3)_{n-1}M_nI_{3n+1}$($R =$ butyl, phenethyl; $M = Pb$, Sn; $n = 1$, 2) toluene/2-propanol mixtures work well as the solvent for the organic salt and the dipping times are relatively short (several seconds to several minutes, depending on the system). Single-phase polycrystalline films of the desired product are achieved using the process, without the need for finding a solvent that can dissolve simultaneously both the organic and inorganic components of the hybrid.

3. Properties

Given the flexibility to tailor independently both the organic and inorganic components of the hybrid, a number of interesting properties have been realized or can be envisioned for crystalline organic–inorganic composites. The inorganic framework, for example, can be used to define a well-ordered polar structure for hyperpolarizable organic cations (see Sect. 1.2), therefore potentially leading to large second harmonic generation (SHG) efficiencies. Several other representative examples are given below.

3.1 Magnetism

The effective dimensionality of an inorganic framework has a fundamental influence upon the observed magnetic properties—especially with regard to the possibility of long-range magnetic ordering. Low-dimensional magnetic systems are approximated in real crystals when the magnetic atoms making up the structure most strongly interact with neighbors that are arranged into clusters (0D), chains (1D), or planes (2D). Given the ability to control the effective dimensionality of crystalline organic–inorganic hybrids by varying the thickness of the inorganic layers and the length of the interleaving organic cations, the hybrid materials offer an ideal opportunity to study low-dimensional magnetism.

Many of the reported magnetic organic–inorganic hybrids are metal halides containing first-row transition metals. The $Cu^{2+}(S = \frac{1}{2})$ cation is a typical example, with a single magnetic orbital (d_{x^2}–d_{y^2}) occupied and essentially an isotropic Heisenberg exchange coupling. For antiferrodistortive CuX_4^{2-} ($X = Cl$ and Br) perovskite sheets (Willett *et al.* 1988), the relatively strong in-plane superexchange interaction is ferromagnetic. The generally weaker interlayer magnetic interaction can be either antiferromagnetic or ferromagnetic, with the magnitude of the interaction being controlled through the length and conformation of the organic cation. While at higher temperature the copper(II) halide perovskites can be considered nearly ideal $S = \frac{1}{2}$ 2D Heisenberg ferromagnets, the weak interlayer interactions lead to an overall 3D magnetic ordering at low temperature. The 3D ordering temperature is in the range 4–70 K for the copper halide structures. Hybrid perovskites containing other metal cations also exhibit cooperative magnetic phenomena, including the Cr^{2+} salts which have in-plane ferromagnetic interactions (like the Cu^{2+} compounds), and the Mn^{2+} and Fe^{2+} salts, with antiferromagnetic coupling between nearest neighbors. Magnetism in the layered perovskites and related low-dimensional systems has been extensively reviewed (see, for example, de Jongh 1990).

3.2 Electrical Conductivity

Electrical measurements on the perovskites $(CH_3(CH_2)_nNH_3)_2MX_4$ and $(NH_3(CH_2)_nNH_3)\ MX_4$ ($M = Cu$ or Cd; $X = Cl$ or Br) indicate that these materials, like most metal halides, are good electrical insulators at low temperature (Tigges *et al.* 1983). At room temperature, a finite but very small conductivity ($\sim 10^{-12}\,(\Omega\,cm)^{-1}$) originates most likely from ionic rather than electronic conduction. In contrast, several tin(II) halide-based perovskites exhibit substantial electronic conductivity associated with the extended inorganic framework (Mitzi *et al.* 1994). In $(C_4H_9NH_3)_2(CH_3NH_3)_{n-1}Sn_nI_{3n+1}$, for example, there is a semiconductor–metal transition as a function of increasing *n* (perovskite sheet thickness). The $n = 1$ compound, $(C_4H_9NH_3)_2SnI_4$, is a semiconductor, with a room-temperature resistivity (pressed pellet sample) of $\sim 10^5\,\Omega\,cm$. The $n \to \infty$ compound, $CH_3NH_3SnI_3$, is a low carrier density p-type metal with a carrier density of approximately 10^{19} holes cm^{-3}, a room-temperature resistivity of $10^{-2}\,\Omega\,cm$, and a Hall mobility of $\sim 50\,cm^2\ (V\ s)^{-1}$. Consequently, the metal halide-based composites exhibit a range of electrical properties (from insulating to metallic) and can provide a framework with substantial electrical mobility.

3.3 Luminescence

Some crystalline composites exhibit strong luminescence that can arise either from the inorganic or the organic component of the structure. For hybrid perovskites containing optically inert organic cations (i.e., those that do not absorb or emit in the visible range), a strong absorption and photoluminescence (PL) peak in the spectra for germanium(II), tin(II), or lead(II) halide systems arises from the radiative decay of relatively mobile excitons associated with the bandgap of the inorganic framework. The position of the PL peak can be tuned across the entire visible spectrum by making substitutions on the metal or halogen site of the perovskite or by changing the thickness of the perovskite sheets (Papavassiliou and Koutselas 1995). Low-temperature ($T < 200\,\mathrm{K}$) electroluminescence (EL) has also been observed in these systems (Hong *et al.* 1992). However, the EL is quenched at room temperature as a result of thermal ionization of the excitons. Photoluminescence from the inorganic framework also occurs in some transition-metal (e.g., Mn^{2+}) and rare-earth metal (e.g., Eu^{2+}) halide perovskites. In these systems, the relevant transitions arise from more localized *d* and *f* band transitions and are therefore more sensitive to the local coordination of the metal cation (e.g., through crystal field splitting) than to the global features of the inorganic framework (e.g., the thickness of the perovskite sheets). In the Mn^{2+} halides, for example, red luminescence arises in manganese(II) chloride-based perovskites with octahedral Mn^{2+} coordination, while strong green luminescence is observed for organic–inorganic hybrids with tetrahedral Mn^{2+} coordination.

Chromophores have also been substituted for the optically inert organic cations within the perovskite framework and have been shown to have an important impact on the optical properties of the hybrid. A series of $(RNH_3)_2PbCl_4$ perovskites (R = 2-phenylethyl, 2-methylnaphthyl, and 2-methylanthryl) has been synthesized, for example, with the goal of varying the energies of the organic molecule singlet and triplet states relative to the exciton band of the inorganic framework (Braun *et al.* 1999). For R = 2-phenylethyl, the organic singlet and triplet transitions are at a higher energy than the inorganic exciton band transition, leading to emission only from the exciton state. For R = 2-methylnaphthyl, the exciton band transition lies between the singlet, S_1, and triplet, T_1, transitions and phosphorescence from the organic triplet state is observed. Finally for R = 2-methylanthryl,the inorganic exciton transition is at a higher energy than both the singlet and triplet transitions in the chromophore, leading to singlet emission from the organic layer. These observations are explained in terms of energy transfer between the organic and inorganic components of the structure.

Similarly, in the perovskite system (AEQT)PbX_4(AEQT = 5,5‴-bis(aminoethyl)-2,2′:5′,2″:5″,2‴-quaterthiophene; X = Cl, Br, and I) the absorption spectra for the hybrids exhibit both the exciton transition from the inorganic sheets and a broader absorption associated with the quaterthiophene chromophore (Mitzi *et al.* 1999a). In the PL spectrum, however, the exciton peak from the inorganic framework is quenched. While there is strong PL from the AEQT molecule for X = Cl, this luminescence is substantially diminished across the series from X = Br to X = I. The PL quenching is associated with a progressively smaller inorganic layer bandgap relative to the chromophore HOMO–LUMO gap and the subsequent energy and/or charge transfer between the two components. It is interesting to note that, as for dye-sensitized inorganic semiconducting surfaces (Sakata *et al.* 1990), charge transfer between the organic and inorganic layers of hybrid crystalline composites can be expected.

3.4 *Potential Applications*

The potentially useful properties and simple thin-film processing techniques described above suggest that organic–inorganic hybrids may prove to be valuable for various technological applications. For example, organic systems have generated substantial interest as semiconducting channels in field-effect transistors (FETs) because of the possibility of simple processing. Spin-coated organic FET channel layers have been demonstrated (Herwig and Müllen 1999) with mobilities as high as $0\cdot1\,\mathrm{cm}^2\,(\mathrm{V\,s})^{-1}$ (comparable to that of amorphous silicon). If larger mobilities were achievable, higher-speed devices and therefore a more extensive range of potential applications could be considered. For the organic–inorganic hybrids, similar thin-film processing techniques can be employed as for the organic materials, but higher electrical mobilities are possible because of the stronger ionic and covalent forces binding the inorganic framework. An FET containing a spin-coated film of $(C_6H_5(CH_2)_2NH_3)_2SnI_4$ has recently been demonstrated with a field effect mobility of $0\cdot6\,\mathrm{cm}^2\,(\mathrm{V/s})^{-1}$ (Kagan *et al.* 1999). These early organic–inorganic FETs currently have the highest reported mobility values for spin-coatable films and higher values can be envisioned as other inorganic frameworks are explored.

A second area of interest involves organic light-emitting devices (OLEDs). Organic materials have the advantages of high luminous efficiency, spectral versatility, and easy processing. However, the organic dye materials are generally insulating and less thermally stable than inorganic compounds. Joule heating during device operation, as well as the necessarily very thin profile of the organic films, leads to device degradation and short lifetimes. Organic–inorganic hybrids offer the possibility of combining the strong luminous efficiency of organic dye materials with the electrical and thermal/mechanical properties of inorganic compounds. A LED based on the organic–inorganic perovskite (AEQT)$PbCl_4$ has been demonstrated (Chondroudis and Mitzi 1999) which has a thick emission layer (3000 Å) and relatively efficient room-temperature EL. This is in contrast to previous work on devices containing organic–inorganic perovskites with optically inert organic cations, which only exhibited EL at low temperature (Era *et al.* 1994).

While the field of electronic devices based on self-assembling organic–inorganic composites is relatively new, the already demonstrated organic–inorganic FET and LED devices offer several interesting opportunities. Other possibilities include the use of the composites in nonlinear optical devices and solar cells. The fact that the devices can be processed using low-temperature techniques, such as spin coating, inkjet printing, and stamping, enables processing on flexible plastic sheets, rather than on heavier, brittle, more expensive inorganic (e.g., silicon or glass) substrates, which are necessary for higher-temperature processing. Hence, organic–inorganic electronics technologies may help to enable flexible, lightweight, low-cost electronic devices and displays.

4. Conclusion

Crystalline organic–inorganic hybrids combine, on a molecular level, an extended inorganic framework and a network of organic molecules. The organic component can interact with the inorganic framework through a range of interactions, including van der Waals, ionic, and covalent forces, and can be used to define the dimensionality of and the interactions between the components of the inorganic framework. Currently, more complex organic molecules are being designed to fit within the composites, with the goal of creating more functional organic layers. This effort includes designing molecules that are strongly luminescent, exhibit nonlinear optical properties, or are photosensitive or electrically conducting. The inorganic framework, which can consist of 0D–3D arrays of corner-sharing, edge-sharing, or face-sharing inorganic polyhedra, may also provide distinct functionality for the material. By controlling the inorganic framework bandgap relative to the organic energy levels, self-assembling multiquantum well structures can be formed and engineered. In addition, the use of the inorganic framework as a template for performing chemical reactions on the tethered organic molecules has begun to be explored and offers a number of interesting possibilities.

While organic–inorganic composites may contain a variety of inorganic frameworks, metal halide frameworks are particularly interesting because of their solubility in a wide range of common aqueous and organic solvents, as well as their relatively low volatilization temperature. These characteristics enable crystal growth and thin-film deposition using a number of simple techniques. In addition, metal halides (e.g., the hybrid perovskites) offer a range of interesting properties including magnetism, high electrical mobility, and strong luminescence. The ability to conveniently process the hybrids, coupled with the potentially useful properties, suggests that organic–inorganic hybrids may find opportunities within a range of electronic and optical applications.

Bibliography

Arend H, Huber W, Mischgofsky F H, Richter-van Leeuwen G K 1978 Layered perovskites of the $(C_nH_{2n+1}NH_3)_2MX_4$ and $NH_3(CH_2)_mNH_3MX_4$ families with M = Cd, Cu, Fe, Mn, or Pd and X = Cl or Br: Importance, solubilites and simple growth techniques. *J. Cryst. Growth* **43**, 213–23

Braun M, Tuffentsammer W, Wachtel H, Wolf H C 1999 Tailoring of the energy levels in lead chloride-based layered perovskites and energy transfer between the organic and inorganic planes. *Chem. Phys. Lett.* **303**, 157–64

Chondroudis K, Mitzi D B 1999 Electroluminescence from an organic–inorganic perovskite incorporating a quaterthiophene dye within lead halide perovskite layers. *Chem. Mater.* **11**, 3028–30

Corradi A B, Bruni S, Cariati F, Ferrari A M, Saccani A, Sandrolini F, Sgarabotto P 1997 Organic/inorganic composite materials: synthesis and properties of one-dimensional polymeric haloplumbate(II) systems. *Inorg. Chim. Acta* **254**, 137–43

Day P, Ledsham R D 1982 Organic–inorganic molecular composites as possible low-dimensional conductors: photopolymerization of organic moieties intercalated in inorganic layer compounds. *Mol. Cryst. Liq. Cryst.* **86**, 163–74

de Jongh L J 1990 *Magnetic Properties of Layered Transition Metal Compounds*. Kluwer, Dordrecht, The Netherlands

Era M, Hattori T, Taira T, Tsutsui T 1997 Self-organized growth of PbI-based layered perovskite quantum well by dual-source vapor deposition. *Chem. Mater.* **9**, 8–10

Era M, Morimoto S, Tsutsui T, Saito S 1994 Organic–inorganic heterostructure electroluminescent device using a layered perovskite semiconductor $(C_6H_5C_2H_4NH_3)_2PbI_4$. *Appl. Phys. Lett.* **65**, 676–8

Herwig P T, Müllen K 1999 A soluble pentacene precursor: synthesis, solid-state conversion into pentacene and application in a field-effect transistor. *Adv. Mater.* **11**, 480–3

Hong X, Ishihara T, Nurmikko A V 1992 Photoconductivity and electroluminescence in lead iodide-based natural quantum well structures. *Solid State Commun.* **84**, 657–61

Kagan C R, Mitzi D B, Dimitrakopoulos C D 1999 Organic–inorganic hybrid materials as semiconducting channels in thin-film field-effect transistors. *Science* **286**, 945–7

Liang K, Mitzi D B, Prikas M T 1998 Synthesis and characterization of organic–inorganic perovskite thin films prepared using a versatile two-step dipping technique. *Chem. Mater.* **10**, 403–11

Mitzi D B 1999 Synthesis, structure, and properties of organic–inorganic perovskites and related materials. *Prog. Inorg. Chem.* **48**, 1–121

Mitzi D B, Chondroudis K, Kagan C R 1999a Design, structure, and properties of organic–inorganic perovskites containing an oligothiophene chromophore. *Inorg. Chem.* **38**, 6246–56

Mitzi D B, Feild C A, Harrison W T A, Guloy A M 1994 Conducting tin halides with a layered organic-based perovskite structure. *Nature* **369**, 467–9

Mitzi D B, Prikas M T, Chondroudis K 1999b Thin film deposition of organic–inorganic hybrid materials using a single source thermal ablation technique. *Chem. Mater.* **11**, 542–4

Papavassiliou G C, Koutselas I B 1995 Structural, optical and related properties of some natural three- and lower dimensional semiconductor systems. *Synth. Met.* **71**, 1713–4

Pecaut J, Le Fur Y, Masse R 1993 Crystal engineering and structural investigations of the 2-amino-5-nitropyridinium salts $C_5H_6N_3O_2^+ \cdot HSO_4^-$ and $C_5H_6N_3O_2^+ \cdot HAsO_4^-$. *Acta Crystallogr., Sect. B.* **49**, 535–41

Petty M C 1996 *Langmuir–Blodgett Films: An Introduction*. Cambridge University Press, New York

Sakata T, Hashimoto K, Hiramoto M 1990 New aspects of electron transfer on semiconductor surfaces: dye-sensitization system. *J. Phys. Chem.* **94**, 3040–5

Sanchez C, Ribot F, Lebeau B 1999 Molecular design of hybrid organic–inorganic nanocomposites synthesized via sol–gel chemistry. *J. Mater. Chem.* **9**, 35–44

Takada J 1995 Organic–inorganic hetero nanosystems as an approach to molecular optoelectronics. *Jpn. J. Appl. Phys.* **34**, 3864–70

Tieke B, Chapuis G 1986 Solid state polymerization of butadienes in layer structures. *Mol. Cryst. Liq. Cryst.* **137**, 101–16

Tigges C P, Drumheller J E, Snively L O 1983 Electrical conductivity measurements of some layered magnetic structures. *J. Appl. Phys.* **54**, 5473–5

Whittingham M S, Jacobson A J 1982 *Intercalation Chemistry*. Academic Press, New York

Willett R, Place H, Middleton M 1988 Crystal structures of three new copper(II) halide layered perovskites: structural, crystallographic, and magnetic correlations. *J. Am. Chem. Soc.* **110**, 8639–50

D. B. Mitzi
IBM T.J. Watson Research Center,
Yorktown Heights, New York, USA

This Page Intentionally Left Blank

P

Paper Products: Classification

Paper and board products are those produced in papermaking and generally involve some converting process. The unit processes of papermaking and their primary functions are as follows:

(i) Refining: fiber structure modification,

(ii) Forming: network formation and water removal,

(iii) Wet pressing: network consolidation and water removal,

(iv) Drying: shrinkage control and water removal,

(v) Size press: surface sizing and coating,

(vi) Calendering and supercalendering: caliper control and surface property modification.

In the papermaking process paper is formed from a uniformly dispersed aqueous suspension of fibers placed onto a screen or forming fabric. Therefore, rice paper (a misnomer—it is actually a veneer cut from the pith of the *Fatsia papyrifera* tree) and papyrus (a woven reed structure produced by ancient Egyptians from which the term paper is derived) do not fall into this category (Hunter 1978). However, specialty products formed as wet-laid nonwoven materials do fall into the category of papermaking. The wet-laid nonwoven process attempts to make textile-like structures. Following web formation the remainder of the process is one of network consolidation, compaction, and water removal through the unit processes of couching, wet pressing, and drying. There are many ways of forming paper and board from the batch process of handmade paper to the special continuous web formers used in the high-speed production of tissue paper. According to Hunter (1978), the first attempts to produce a continuous web of paper were motivated by the strong hold that French handmade paper unions had on paper production.

The raw materials of papermaking are many and varied. The principal components are wood fiber, including softwood and hardwood, fillers, nonwood fibers, and a variety of chemical additives that may be used to improve the efficiency of the papermaking process as well as to enhance the end-use performance of the product. Softwood and hardwood fibers are produced from wood chips by mechanical and/or chemical means. Mechanical pulps include stoneground wood (SG), thermomechanical pulp (TMP), and chemithermomechanical pulp (CTMP). Chemical pulps are mainly produced by the kraft process, although a few sulfite pulp mills are still in operation. Nonwood fibers might include plant and synthetic polymer fibers, as well as glass, ceramic, and metal fibers.

Many paper products require that the product be produced from bleached fibers. In the bleaching process the residual lignin is removed to minimize losses in fiber strength and other chemical components, e.g., hemicelluloses, while at the same time producing a clean, bright pulp. Pulp cleanliness is not only important to the optical properties of paper, but also for electrical-grade papers that are sensitive to impurities and the presence of metal ions. There are many different types of bleaching processes; however, the principal bleaching agent is chlorine, which has been the subject of much criticism because of environmental concerns. This has led to the development of both elemental chlorine-free (ECF) and total chlorine-free (TCF) bleaching systems. Unfortunately TCF-bleached pulps suffer some loss in pulp physical properties including fiber strength and brightness (Heimburger 1998).

Refining, or beating, a term used interchangeably with refining, is part of stock preparation that occurs prior to the sheet-forming process to develop the papermaking potential of the pulp. Balance is sought between the development of paper properties and productivity on the paper machine. Productivity is mainly limited by water removal of which refining has a negative impact. In terms of the paper product aesthetics, converting and end-use performance refining are critical process elements. Refining can modify the structure of the papermaking fiber in a number of different ways, i.e., internally and externally fibrillate, by producing fines, inducing or removing fiber curl and kink, and reducing fiber length (Page 2000). The fiber can be made more flexible and conformable, and its surface area can be increased to promote bonding between the fibers in the network. Fiber shape can also be modified by making the fiber straighter, i.e., removing kinks and curl, or by purposely curling and microcompressing it. The former treatment is designed to produce a paper that is denser, stronger, and stiffer, while the latter treatment could be used to produce a paper that is softer and more bulky. Refining is considered to be an inefficient mechanical process. As refining proceeds there is an increase in water uptake owing to fiber swelling, and the fines or cell wall debris produced adversely affect the drainage behavior of the pulp. Water retention value (WRV) expressed as mass of water per mass of cell wall material is a common measure of water uptake in the cell wall (Anonymous 1991). One common measure of the drainage behavior of a pulp is known as Canadian Standard Freeness (CSF) expressed volumetrically in milliliters of filtrate (Anonymous 1998). If used intelligently, CSF can provide useful information, such as indicating the level of pulp refining. An unrefined softwood pulp would have a CSF of about 740 mL, moderately refined of about 500 mL, and well refined of 100 mL

---

* Cross references marked by an asterisk are included in this volume.

# Cross references marked by a hash can be found by consulting the Encyclopedia of Materials: Science and Technology.

Table 1
General converting processes.

Converting process	Comment
Coating	Aqueous, nonaqueous
Calendering and supercalendering	Including soft-calendering
Printing	Impact and nonimpact processes
Forming and molding	Corrugating, egg and fruit cartons, paper plates and cups
Laminating	Combining different sheet materials, metal, plastics
Impregnation	Sizing, strength additives, Formica
Modification of deformation behavior	Sack kraft, nonwovens
Gluing, bonding, jointing	Tube making
Size reduction	Slitting and rewinding, sheeting, business forms

or less. The initial or unrefined CSF of a hardwood pulp would be $\sim 600\,\mathrm{mL}$.

Converting is one or more processes that convert the large roll of paper from the paper machine, i.e., the machine or Pope reel, into a useable product. Converting processes may involve size reduction and sheeting, the production of smaller well-wound rolls of paper, coating and supercalendering, molding, embossing, laminators, corrugating, and the manufacture of corrugated containers. Table 1 summarizes the main areas of converting. Although some of these processes are specific product related, such as corrugating, they nevertheless may be combined and used to produce less conventional products. For example, corrugating has been used in the production of single-face board (a single liner bonded to corrugating medium) to produce filter-type materials, where the single-face board is wound to form a porous cylinder. It should also be noted that the converting processes given in Table 1 are not exclusively used off the paper machine. Coating, impregnation, calendering, including soft-nip calendering, are examples of converting processes that can be performed on the paper machine.

Paper products are many and varied, and can be classified into one of two broad categories: commodity products (high-volume, relatively low-cost products) and specialty products (essentially the converse). Paper is always in competition with other materials, such as plastics and metals. However, paper can also be combined with other materials to create new products using the converting process of lamination. Predictions have also been made of a significant reduction in paper usage should the paperless society, a product of the information age, ever become a reality. Perhaps the use of paper will decline in some areas; however, it is also possible that this will be offset by growth in new and unforeseen uses. There are not many textbooks dedicated to a comprehensive overview of paper and board products, although there is much information in patents and paper-related journals. However, Paulapuro (2000) is highly recommended. Although written from a Scandinavian and European perspective, it adequately covers the North American scene.

1. Commodity Paper and Board Products

The major high-volume, low-cost products include tissue, newsprint, printing and writing papers, linerboard, medium, and construction grades. The categorization of paper and board is somewhat arbitrary with respect to basis weight (mass per unit area). We might expect paperboard to coincide with some level of basis weight for which there are many measures, some being product specific. Table 2 gives the range of production speeds (approximately maximum) and basis weights associated with a range of commodity products. Also given in Table 2 is the relative importance of a number of physical properties associated with converting and end-use performance.

1.1 Containerboard (Linerboard and Corrugating Medium)

The properties associated with paper and board products are many and varied. TAPPI and ASTM test methods total about 500. The properties of paper and board can be conveniently grouped into the categories given in Table 3.

1.2 Tissue

Tissue and toweling manufacture is a highly specialized and competitive branch of papermaking, and thus very proprietary. Tissue is one of the lowest basis weight products produced, typically in the range $10\text{–}20\,\mathrm{g\,m^{-2}}$. The furnish is produced from special market pulps, and some production facilities are

Table 2
Paper products–property matrix.

Product	Speed ($\mathrm{m\,s^{-1}}$)	Basis weight ($\mathrm{g\,m^{-2}}$)	Mechanical	Optical	Dimensional stability	Printing
Tissue	38	15	X	X		
Newsprint	27	45	XX	XXX	XX	XXX
Fine papers	20	70	X	XX	XXX	XXX
Corrugating medium	18	125	XX		X	
Linerboard	14	205	XX	X	X	XX

Table 3
Test methods for paper and board products.

Category	Examples of properties
Structure and composition	Basis weight, caliper, apparent density, fiber orientation, formation
Mechanical properties	Elastic properties, tensile, compressive and shear properties, burst and tear properties
Physical properties	Optical properties including color, thermal conductivity, electrical properties
Surface properties	Smoothness, friction, surface strength, linting
Porous nature of paper	Void volume, pore size distribution, wettability, surface energy
Paper durability	Fold endurance, natural and accelerated aging tests
Converting and end use	Printability, print quality, runnability
Aesthetics and subjective properties	Appearance and uniformity, printing performance, tissue quality, softness

based on 100% recycled fiber. The furnish is beaten to some intermediate level, e.g., 400 mL CSF. Generally tissue is made on a high-speed former at 25–30 $\mathrm{m\,s^{-1}}$. The web is supported through the consolidation and wet pressing stages before being transferred to a Yankee dryer. The Yankee dryer consists of a large-diameter drying cylinder with a highly polished surface. The web adheres to the surface of the dryer, which may be enhanced by the use of additives and results in more efficient drying of the web. A creping doctor is used to separate the web from the surface of the dryer to which it is bonded. In addition to separation, the most important function of the doctor is to change the properties of the web dramatically. The softness of the tissue is derived substantially from doctoring. The bonds between the fibers are partially broken and the fibers are both curled and microcompressed by the doctor blade in the process of web detachment. Special converting machinery is used to produce toweling, which consists of many layers of tissue, as well as the ubiquitous toilet roll.

1.3 Newsprint

Newspapers are produced from rolls of newsprint having a basis weight of about 45 $\mathrm{g\,m^{-2}}$. The basic newsprint furnish comprises mechanical pulp and chemical pulp. The high-yield mechanical pulp is important in controlling the optical (particularly opacity) and printing properties of the sheet. The chemical pulp, or reinforcing pulp to which it is often referred, is important for the runnability of the web both on and off the paper machine. Runnability refers to the ability to get the web through the paper machine or printing press with the minimum of web breaks, ideally none. Newsprint is produced using twin wire formers to maximize productivity and to minimize two-sidedness. In fact some publishers insist in their specifications that newspapers be printed on paper produced on a twin wire former.

1.4 Printing and Writing Papers

Printing and writing papers can be both commodity and specialty products. The basis weight range can vary considerably and includes both uncoated and coated papers. As shown in Table 2, mechanical, optical, dimensional stability, and printing properties are very important. These papers are produced in both roll and sheet form for roll- and sheet-fed printing processes. The printing process, end-use performance, and economics dictate composition, coating, and the finish required. Printing includes impact and

Table 4
Coated paper grades.

Grade	Brightness	Gloss
Matte—premium	92–94	15–20
Matte—standard	82–84	10–15
Dull no. 1	85–87	20–30
Velvet no. 2	84–86	20–30
Velvet no. 3	82–84	25–35
Velvet no. 4	80–82	25–35
Enamel no. 1	85–87	65–70
Enamel no. 2	84–86	62–65
Enamel no. 3	82–84	57–62
Enamel no. 4	80–82	50–55

nonimpact processes. The former includes gravure (intaglio), offset (planographic), and letterpress (relief). Letterpress is one of the earliest printing methods now being replaced by flexography where relief is created using polymeric rather than metallic plates. Common nonimpact printing processes are inkjet printing and electrophotography.

Printing and writing papers are generally specified with respect to composition, brightness, and gloss (Table 4). Paper is usually coated to achieve high levels of gloss through the use of supercalendering or soft-nip calendering. Paper grades are referred to as wood- free or merchant grades, i.e., those not containing mechanical pulps, while the term publication grade is used for those containing a combination of bleached chemical and mechanical pulp fibers. Printing and writing papers contain high levels of hardwood fiber to improve formation and surface characteristics. Mechanical pulp fibers used in publication grades serve to increase the sheet opacity.

In addition to optical properties, offset papers require high levels of surface strength to resist picking and surface disruption when high-tack printing inks are used. Dimensional stability (water resistance) is also important with respect to register when two or more colors are being printed. Gravure requires a very smooth paper surface to provide intimate contact with the gravure printing cylinder to facilitate ink transfer. In general all printing papers must have a balance between ink absorbency and ink hold out. This is especially dependent on the pore structure and wetting characteristics of the surface. In addition to enhancing optical properties, coating modifies the pore structure of the surface to control ink absorbency and hold out. Thus coating enhances print quality, a highly subjective characteristic.

The long-term survival of paper or its archival value is a major concern for libraries and others involved in information storage. For example, the US Library of Congress has about three million books in a tender or advanced state of degradation. Earlier, it was not realized that acid-based papermaking ($pH < 6$) would greatly reduce the life of paper from acid hydrolysis. There is now a much greater use of alkaline- as opposed to acid-based papers in the production of books. Alkaline papermaking has many advantages with few drawbacks. A calcium carbonate filler can provide a reserve that preserves the alkaline nature of the paper.

It is interesting that early handmade papers such as those made in Europe from the thirteenth to seventeenth centuries had excellent longevity and some are still in nearly mint condition. However, the inadvertent use of alum to improve the effectiveness of gelatin sizing caused some of these papers to be in very poor condition from acid-papermaking conditions (Waterhouse and Barrett 1991).

1.5 Containerboard

Corrugated containers (cardboard boxes) are manufactured from linerboard and corrugating medium collectively referred to as containerboard. Corrugating is the major converting process in a box plant that produces corrugated board blanks. These are subsequently processed by “flexo-folder-gluers” to produce printed corrugated containers in the flat or “knocked down” condition. Corrugated containers are engineered structures that protect the contents during shipment and storage, and in some instances need to be converted into a point-of-sale display item to promote the sale of the contents with appealing graphics.

In addition to the printing requirements, strength is also important. This includes top to bottom or stacking strength, the ability to resist creep particularly in cyclic environments, and end-to-end loads that might be caused by forklift trucks. Containers are often tested to determine their ability to withstand impact loads as well as vibration. The stacking strength of a corrugated container is mainly dependent on the ability of the walls of the container to withstand compressive loads, their flexural stiffness, as well as container geometry. For US interstate shipping and commerce, strength specifications are contained in Rule 41 Item 222. These properties are specified for a given container geometry and size by the compressive strength of the combined board, or the burst strength of the combined board and basis weights of its linerboard facings. Examples for single-wall corrugated fiberboard boxes are given in Table 5.

The row of data in bold print in Table 5 relates to a common container for a maximum load of 290 N, and having a typical construction of two 42 linerboards with possibly a $127\,\mathrm{kg\,m^{-2}}$ medium, with a burst strength of 1.4 MPa. An alternative specification (right-hand column of Table 5) gives a compressive strength, known as the edgewise compressive strength, of $5.6\,\mathrm{kNm^{-1}}$ for the combined board. This alternative specification could be for a

Table 5
Single-wall corrugated fiberboard boxes.

Maximum mass of box (kg)	Maximum outside dimensions (L + W + H) (m)	Minimum burst test (MPa)	Minimum combined mass of facings ($\mathrm{kg\,m^{-2}}$)	Minimum edgewise compressive strength ($\mathrm{k\,N\,m^{-1}}$)
9	1	0.86	0.25	4.0
16	1.3	1.03	0.32	4.5
23	1.5	1.21	0.37	5.1
30	**1.9**	**1.38**	**0.41**	**5.6**
36	2.1	1.72	0.54	7.0
43	2.4	1.90	0.67	7.7
54	2.7	2.42	0.88	9.6

wall construction that is a lower basis weight than one based on burst strength, providing the manufacturer with an opportunity for raw material savings. Edgewise compressive strength of the combined board is dependent on the cross-machine direction compressive strength of its component linerboards and corrugating medium. Only about 60–85% of the compressive strength of the components is realized.

Linerboard is manufactured from virgin high-yield unbleached kraft pulps and recycled pulps including old corrugated containers (OCC) and box plant waste or clippings. Most linerboard is manufactured with a base sheet from a higher yield pulp than that used to produce the secondary or top layer of the sheet. The main purpose of the top layer is to provide a good printing surface. Therefore, more refining is used for the secondary furnish, 300 mL CSF as opposed to the base sheet of typically 650 mL CSF. Since CD compressive strength is of paramount importance the paper maker attempts to make as "square" a sheet as possible, MD/CD = 1.

Corrugating medium is manufactured from a variety of furnishes including neutral sulfite semichemical (NSSC) and OCC. In a common board construction of 205/127/205 $\mathrm{g\,m^{-2}}$ the medium represents approximately 30% of the total basis weight. The medium also contributes to the CD compressive strength and plays an important role in resisting crushing where MD compressive strength is important.

2. Specialty Papers

Paper products, particularly specialty papers, are many and varied. Listings of paper products and descriptions can be found elsewhere (Anonymous 1999, Paulapuro 2000, Gould 1976, Mosher and Davis 1968, Casey 1983, and Lavigne 1993). A general listing of paper products is given in Appendix 1. Space does not permit a detailed description of all of the products listed, including raw materials, papermaking process, and converting and end-use requirements, but nevertheless it illustrates the diversity of products in papermaking.

2.1 Playing Card Stock

Playing cards, used both for entertainment, games, and educational purposes, have a long history originating in China. The modern playing card has high demands on its manufacturing and end-use performance. It is a coated laminated board product with a high level of print quality. In addition to the need to be dimensionally stable and warp free, the board must maintain a high level of mechanical performance during its usage, referred to as "snap," the ability of the card to maintain its flexural and elastic stiffness during repeated shuffling and dealing.

2.2 Filters

There are many different types of filters with a multitude of end uses ranging from teabags and vacuum cleaner bags to automobile air cleaners. The general but conflicting requirements of a filter are high permeability but a small pore size. Wood, plant, and glass fibers have long been used in filter materials, and also synthetic fibers, particularly as nonwovens. Although synthetic fibers are more costly they have a geometric advantage in control of their cross-sectional shape. Fiber types include polyester, rayon, nylon, Nomex, and Kevlar. The overall efficiency of the filter in terms of particulate removal is inversely proportional to the square of the fiber diameter. Therefore, small-diameter fibers are highly desirable, commensurate with their ability to achieve the desired level of permeability.

2.3 Shoe Innersoles; Luggage

Many specialty products are basically polymer-impregnated fiber networks manufactured using both softwood and hardwood fibers, sometimes with synthetic fiber additions. There are three main methods of polymer addition: furnish or beater addition, wet web saturation, and dry web saturation. In these methods, the polymer may be added before web formation, to the wet web after initial consolidation, or to the dry web either on or off the paper machine, respectively. The polymers available for web impregnation (natural rubber, neoprene, acrylics, and styrene butadiene) have a range of glass transition temperatures. Polymer addition can significantly improve the properties of the fibrous network, including improvement of interfiber bonding, abrasion resistance, fold endurance, and tear and tensile strength.

Unknown to many, footwear has a foundation innersole made from a polymer-impregnated fiber network. The innersole or shoe board material is made on an inclined Fourdrinier paper machine that operates at less than $0.5\,\mathrm{m\,s^{-1}}$. Wet web saturation is used for polymer addition (Casey 1983). A wide range of products is available to satisfy a variety of shoe, boot, and sneaker styles, and methods of fabrication. Shoe board is essentially a substitute for leather as are other substitutes such as in luggage. In luggage applications the board material would be precoated, coated, and embossed to produce a "leather-like" finish.

2.4 Electrical-grade Papers

Examples of electrical-grade papers include electrolytic- and electrostatic-grade capacitor papers, cable and wire insulation, transformer insulation, and battery separators. Electrical-grade papers can be manufactured from both unbleached and bleached grades of pulp; however, purity is a key factor, e.g., absence of metallic compounds and chlorides, in obtaining the required insulating and dielectric performance. Electrostatic-type capacitor-grade paper is manufactured using deionized water. The well-beaten stock produces a paper that is nearly glassine and must be free of pinholes.

Transformer board is in the high basis weight range of $500–3000\,\mathrm{g\,m^{-2}}$. Water and pulp quality are also important requirements for this product. The papermaking process is an intermittent/continuous one. The board is produced by winding a wet web onto a forming cylinder, and when the desired basis weight is obtained the multilayer wet web is released from the forming cylinder for subsequent consolidation, drying, and calendering processes. The apparent density of the final product can be in the intermediate to high range.

Separators for both primary and secondary batteries have long been produced in the papermaking process. In lead–acid batteries, a paper-based separator, which replaced the earlier use of wooden veneers, must physically separate the negative and positive plates. The separator should have, in conjunction with the electrolyte, a low ohmic resistance, important for battery performance at low temperatures. The filter functions of the separator require that pores are sufficiently small and tortuous to prevent dendritic bridging. Other requirements include longevity in the harsh environment of battery acid. In the earlier wood veneer separators, cellulose was protected from acid hydrolysis by lignin. For cellulose fiber separators, this protection was provided by phenolic resin that was added using the wet web saturation process. However, with the advent of the maintenance-free battery in the mid-1970s, the phenolic–cellulose separator was replaced by a noncellulose glass mat.

2.5 Decorative Laminates

It is not often realized that bathroom and kitchen counter tops are paper-based products of laminates that evolved in the early 1900s. Saturating kraft is the basis of this product with the paper made on a Fourdrinier-type paper machine from very free-draining unbleached softwood and hardwood kraft pulps. The critical structure of the paper is such that it will have sufficient void volume and be readily wettable by the resin. The saturating kraft paper is subsequently impregnated with phenolic and melamine formaldehyde resins. A number of plies or layers are used to make the required laminate and are bonded together using heat and pressure. A decorative printed layer is then bonded to the outer surface to provide the necessary aesthetics together with a protective or overlay sheet.

2.6 Currency Papers

Paper money has been around for a long time and is expected to continue despite the wide use of private and business electronic transactions. The end-use requirements of this product are particularly demanding. The ability to maintain its performance over a reasonably long period of time in circulation is highly desirable. The production of currency paper is a very specialized papermaking process. Currency paper is generally made from a well-beaten mixture of linen and cotton fibers. Wet strength, a major end-use requirement, requires addition of wet strength agents. Whereas untreated paper retains only about 4–5% of its dry strength when wet, currency paper requires a strength retention of 40–50%. Melamine–formaldehyde has been the wet strength resin of choice, but environmental concerns about formaldehyde has promoted the use of alternative resins.

It is well known that currency paper loses its crispness during circulation. There is also the necessity of maintaining flexural stiffness so that automatic teller equipment can operate correctly. Therefore, with improvements in furnish selection and treatment, and resin addition, the useful life of paper currency can be extended.

2.7 Metallized Papers

Joining paper with other materials such as plastics and metals can take many different forms. Paper metallizing is one such important process, involving application of a thin layer of metal to the surface of paper. Labeling is one of the largest end uses for metallized paper including beer and beverage labels, gift-wrap, and wall decorations, using silver, titanium, stainless steel, and aluminum.

The two basic methods of metallizing are direct vacuum metallization and transfer metallization. The main method of metallizing employs a batch or semicontinuous operation where the metal is typically aluminum. Ceramic boat temperatures in the range 1400–1500 °C and a high vacuum of 0.13–1.3 torr are required. The thickness and uniformity of the aluminum layer is controlled by boat temperature, web speed, and the rate of aluminum feed. Operations of unwinding, vaporization, metallizing, and rewinding are all performed under vacuum. Prior to metallizing, the base paper is treated with a lacquer coating, which is vital to promote the adhesion of metal as well as providing curl control. After metallizing, the substrate is remoisturized and may be given a protective topcoat.

Indirect or transfer metallization consists of metallizing a plastic film, applying an adhesive to it, and after partial drying, laminating it to the base paper. After curing, the plastic film is stripped for reuse. In this process, the base paper can be rougher and more porous without affecting the reflective properties of the metal film and does not have to be dried and remoisturized.

2.8 Paper Tubes

There are many varieties of tubes ranging from transparent tape dispensers to paper tube molds for forming concrete columns. The two main methods of tube making are spiral and convolute winding. The spiral method is one of continuous tube production where tube length is determined by a cutoff saw. The paper tube mandrel or former is stationary, and the tube is formed by winding several layers of paper strips spread with adhesive. After cutting, the tubes are dried in an oven. Paper rolls for tubes are generally made from recycled paper.

For tubes for pouring concrete columns, the design specifications are quite challenging in that the tubes must not fail during both the pouring and curing processes. A main performance criterion is the ability of the tube to withstand a large hydrostatic pressure, with a minimum of paper to achieve the optimum strength to weight ratio. This permits the tube manufacturer an opportunity to be cost competitive. There also must be a reasonable level of wet strength performance and dimensional stability. Tube testing is important to validate performance such as burst strength; however, testing of very large tubes is difficult and the performance is often extrapolated from that of smaller tubes.

2.9 Vulcanized Board

Vulcanized board is an interesting specialty product in that the base paper undergoes significant chemical and structural alteration. The resulting board is extremely tough in both in- and out-of-plane properties. The basic process involves treating the base paper with zinc chloride or sulfuric acid, causing a partial surface conversion of cellulose I to cellulose II. As a consequence, the board is densified during consolidation of the structural multilayers, with an improvement in interfiber bonding. As a result of the partial conversion of the cellulose, the board becomes more hydrophilic. Changes in fiber structure such as curl reduction, fibril orientation, and enhanced interfiber bonding provide high tensile energy absorption and toughness.

2.10 Glassine

Glassine is used in a wide variety of applications including cereal and other food packaging. It is produced from a well-refined furnish where fibers undergo extreme hydration. Improvements in the process to allow some reduction in refining have been brought about by size press treatments. The furnish has traditionally included both high-yield softwood kraft pulps and hardwood sulfite pulps. Headbox temperatures are usually quite high, about 60 °C, to promote drainage of the low CSF stock. The density of glassine, depending on the specific product, is typically very high, achieved by supercalendering the moist sheet. Supercalendering, a process where the sheet is passed between a hard steel roll and a relatively soft roll that creates a small, extended nip, may consist of a series of up to 16 nips. The final sheet has an apparent density in excess of 1000 $\mathrm{kg\,m^{-3}}$ (cell wall density of cellulose is approximately 1550 $\mathrm{kg\,m^{-3}}$), and therefore has low air permeability as well as being grease resistant. A highly desirable attribute is the absence of pinholes, which are tested for using dye penetration.

2.11 Embossing Papers

Papers to be embossed need a high level of extensibility (elongation to breakage) to avoid rupture during embossing. Controlled moisture and temperature can also ensure satisfactory embossing. In addition, the paper must have a level of permanent set to maintain the embossed pattern; fortunately, paper is viscoelastic–plastic and has this characteristic. The elongation or molding characteristics depend on the types of fiber, their treatment, and the drying restraints applied.

2.12 Carbonless Papers

Before the advent of high-speed office copiers, duplication was carried out using a "carbon," paper to which a wax-based, carbon-containing film had been applied. In 1954 the National Cash Register Company began marketing the carbonless paper, which has undergone extensive development to improve and extend its application. Carbonless papers have categories of chemical and mechanical transfer, although chemical copy papers are preferred. The chemical mated system uses matching sheets to create an image in response to pressure. The back of the top sheet (CB) has a coating applied that contains microcapsules (produced by the process of microencapsulation) of a colorless organic dyestuff. This is matched to a second coating on the front of the bottom sheet (CF), acting as a receptor for the dye released from the microcapsules when they are ruptured by impact, such as typing or pressure from a ballpoint pen. The sheets between the CB and CF in multipart forms are coated on their front side with the receptive coating and on their backside with the color microcapsules. The microcapsules are formed by coacervation, interfacial polycondensation, and solvent-induced processes. Coacervation consists of emulsification of the liquid dye in a dispersion of the encapsulating material. The solvent characteristics are then changed to separate the dye from the wall material, which is subsequently polymerized. The resulting microcapsules are about 1–10 μm in diameter. Coating of the base paper with microcapsules is a critical operation and starch is typically used as a spacer particle to prevent damage during winding and other converting operations. The base paper must have excellent formation, and sizing agents may be used to achieve adequate strength and air and solvent resistance.

2.13 Ceramic Papers

The papermaking process can be used to produce papers with high levels of inorganic fiber, metals, and synthetic fibers. Ceramic papers made in this way are precursors to subsequent converting processes to form and mold the material as well as adding high-temperature bonding agents, if not already incorporated. In comparison with wood fibers, ceramic fibers are stiff, noncomformable, and devoid of any natural bonding agent. One approach to forming ceramic papers in papermaking is to incorporate a small fraction of wood fiber with low-level refining and possibly polymeric bonding agents.

The ceramic green sheet, to which it is sometimes referred, has been proposed for use in catalytic converter applications and other high-temperature gas treatment processes. In the catalytic converter, the ceramic green sheet after corrugating can be rolled into a porous tube and fired to produce a true ceramic. Cellulose and other organic additives are pyrolyzed during this process. Following this operation a platinum film can be applied and fired.

2.14 Dissolving or Water-soluble Paper

Some specialties are produced in very limited quantities since their demand is quite low. Water-soluble paper was at one time produced in the USA, but is now manufactured in Japan. It is remarkable because it "dissolves" in a fraction of a second when immersed in water, although it is quite strong in the dry state. Dissolving is a misnomer since the cellulose does not go into solution, but swells at an extremely rapid rate, immediately breaking the bonds between the fibers in the network.

Applications have included in espionage, gardening, and novelty applications. The technology for producing such a paper closely parallels that used to produce superabsorbents, such as used in feminine hygiene products.

2.15 Xerographic Paper

Xerographic paper is important in both printing and copying. The development of high-speed color printers and copiers has placed even greater demands on runnability and printability. Fused toner particles, which have been deposited on the surface of the paper by electrostatic attraction, produce the printed image. The quality of the printed image is dependent on both the physical properties of the toner and paper as well as the mechanical condition of the printer or copier. The quality of the printed image is dependent on the optical, formation, and surface characteristics of the paper. Furthermore, the uniformity of charge on the surface of the paper is also likely to affect print uniformity. Dimensional stability, i.e., resistance to curl formation induced by heat and/or moisture changes, is probably one of the most demanding physical characteristics of xerographic-type papers.

3. Conclusions

This article attempts to provide a broad overview of paper products, including the raw materials, papermaking, and converting processed involved. The interactions of these basic elements are vital in controlling the productivity, properties, and end-use performance of the product, and importantly its cost. Converting is considered to include nine basic processes each one of which may have many subcategories. Some of these can be performed on the paper machine, while others are off-machine processes. Paper products have been divided into two categories, namely commodity and specialty. The latter category is meant to include those products that are relatively low volume but have high added value. The wide range of specialty paper products discussed indicates the tremendous variety and novelty involved in papermaking.

See also: Papermaking*; Paper Products: Container Board#; Paper Products: Food Packages#; Paper Products: Security Applications#

Appendix 1: Alphabetical Listing of Paper and Board Products

Abrasive backing
Abrasive papers
Absorbent papers
Acoustical board and tile
Adding machine paper
Adding machine rolls
Album paper
Albums
Alkaline-based printing papers
Announcements: commercial, social, wedding
Antitarnish papers and board
Art papers
Asbestos paper and board
Aseptic (extended shelf life) paperboard and packaging
Asphalted paper and board
Bag papers: combination rope and kraft shipping sack, recycled bag and sack, unbleached kraft bag and sack, unbleached kraft shipping sack, other bag and sack
Bags (materials of construction): cellophane, glassine, laminated textile, parchment, waterproof, waxed, wet strength
Bags: multiwall shipping sacks, refuse sacks
Battery paper and board
Beaming paper
Bible paper
Binders board
Black paper
Blanks: coated, plain, mill
Blotting: plain, coated, embossed, enameled, tablet
Bond: chemical wood pulp, cotton (rag content), recycled content, No. 4 cut size (plain copier paper)
Book, alkaline based (acid free, neutral): coated, uncoated
Book coated: cast coated, machine coated, dull C2S, C1S, lightweight (lwc)
Book uncoated: antique, bulking, eggshell, English finish, machine finish, supercalendered
Book offset: coated, uncoated, web
Book end leaf papers
Boxboard folding and set up (14 grades)
Box cover papers (5 grades)
Box liner paper
Boxes, corrugated
Boxes, solid fiber
Boxes, folding (14 grades)
Boxes, set-up (rigid) (10 grades)
Braille paper
Bristols (10 grades)
Burlap lined paper
Business forms
Butchers papers
Button board
Calendar papers
Calendars
Can board stock
Cans (paper, composite)
Cap board
Caps and closures (4 grades)
Carbonizing papers
Carbonless copy base stock
Carbonless copy papers
Carbon paper (4 grades)
Cards, cut
Car liner
Catalog
Cellophane
Ceramic papers
Chart paper (electrosensitive)
Check paper
Chocolate dividers
Cigarette papers
Cloth-lined papers
Coasters
Coating base stock: groundwood papers, book papers
Computer printout paper
Concrete moulds
Condenser paper: electrolytic, electrostatic
Construction (colored school) paper
Copier papers: bond no. 4, inkjet, laser, reprographic
Core stock
Cores
Corrugating medium: semichemical, recycled
Cover papers (10 grades)
Covers and binders
Creped papers
Cup stock
Cups, baking
Cups, drinking (5 grades)
Currency bands and coin wrappers
Currency paper
Decorative laminates
Decorative papers
Dishboard
Disposable garments (14 products)

Doilies
Drawing papers
Dusting paper
Dynamite shell, blasting paper
Egg cases
Electrical transformer board
Electrical cable paper
Electrical insulation paper
Embossed boards and paper
Embossing papers
Envelope stock (4 grades); envelopes (23 types)
Extensible paper
Facsimile paper (bond no. 4 reprographic papers)
Felt, dry (roofing, saturating, organic glass mat)
Fiberboard
Fiber drums
File folder stock
Filter paper
Filters
Fish paper
Flameproof paper and board
Flexible packaging
Fluorescent papers
Food board
Friction board
Gaskets
Gift-wrapping
Glass papers
Glassine (plain, printed, coated, embossed, laminated, metallized, wet strength, waxed)
Greaseproof paper
Grease-resistant paper and board
Greeting card stock
Groundwood printing papers (coated, uncoated, supercalendered, LWC)
Gummed paper (gumming paper)
Gypsum board liner
Hanging paper (wallpaper)
Hardboard
Heat-sealing papers
Hygiene paper specialties (hospital, doctor, dental)
Illustration board
Impregnating paper (absorbent, dry felt, saturating)
Imitation parchment
Industrial packaging and converting papers and board?
Industrial wipes
Inkjet copier papers (bond no. 4 reprographic)
Insecticide papers
Insulating building papers
Insulation paper
Interleaving papers
Ion-exchange paper
Kraft converting paper
Label stock (uncoated and coated)
Laminating stock
Lampshade papers (fire- and flame-retardant, translucency)
Laser copier papers (bond no. 4)
Latex-treated paper and board
Layer board
Layer insulation paper
Ledger (coated, rag content, sulfite, recycled content)
Lens paper
Linen paper (linen finish, paper made from linen rags)
Linerboard (10 grades)
Liners
Lithograph (coated, uncoated)
Manila paper and board
Map paper
Marble papers
Masking papers
Mat board
Matchbook board
Masking tape
Matrix board
Mats (bath, automobile, place)
Metal fiber
Metallic papers and board
Mica papers
Micro paper
Milk carton stock
Mill boards
Mimeograph paper
Mulch paper (garden and agricultural)
Music paper
Napkins
Needle paper
Newsprint (11 types)
Nitrating paper
Noiseless paper
Novelties, toys, and games
OCR paper (optical character recognition)
Oiled papers and board
Onionskin
Opaque
Ovenable paperboard packaging (microwavable)
Overlay papers
Packing board
Packing and cushioning materials (dunnage)
Pallets
Panel board
Pan liners
Papeterie papers
Parchment papers (imitation, vegetable)
Partition stock
Pasted board
Pattern board
Pattern paper
Phenolic impregnated papers
Photographic base paper
Photomount stock
Placemat stock
Plates (molded pulp, pressed paper, baking, ice cream)
Playing card stock
Pleating paper

Polycoated paper and boards
Postcard
Poster board
Poster paper
Pressboard (genuine, imitation)
Pressure-sensitive paper
Proofing paper
Pyroxylin-coated papers
Railroad board (coated, uncoated)
Reinforced papers
Release papers
Reproduction base papers
Reproduction papers (blueprint, blackline, brownprint, diazo, whiteprint, thermal, xerographic, diffusion transfer)
Ribbon papers
Rope papers
Rotogravure papers
Rust inhibitor papers
Sanitary napkins
Saturating paper
School papers
Self-sealing papers
Sheathing (plain, waterproof)
Shelf paper
Shirt board
Shoe board
Shredded paper (cellophane, glassine, packing, parchment, tissue, waxed)
Slipsheets (corrugated, solid fiber)
Stencil board and papers
Straw paper (drinking)
Synthetic fiber papers
Tablet paper
Tabulating card stock
Tag board
Tag stock
Tape, gummed
Tape, pressure sensitive
Teabag paper
Technical specialty papers
Template board
Text papers
Ticket stock
Tissue facial
Tissue papers (9 types)
Tissue toilet (rolls, household, interfolded sheets, jumbo rolls)
Toweling
Tracing papers
Translucents
Transfer paper
Tube board
Tubes
Twisting paper
Tympan paper
Vaporproof paper and board
Vellum papers
Velour papers
Vulcanizing paper
Vulcanized fiber
Wallboard
Wallpaper
Watermarked papers
Waterproof boards
Waterproof building papers
Waterproof papers
Water-soluble paper
Waxed papers (wet-waxed, dry-waxed, wax-laminated, wax-sized)
Wet strength papers
Wrapping papers
Wrapping tissue
Wrappers (13 types)
Writing papers
Xerox copy papers

Bibliography

Anon 1991 Water retention value (WRV) UM 256. In: *TAPPI Useful Methods*. TAPPI Press, Atlanta, GA, pp. 54–6

Anon 1998 Freeness of pulp (Canadian standard method) T227 om-94. In: *1998–1999 TAPPI Test Methods*. TAPPI Press, Atlanta, GA

Anon 1999 *Lockwood-Post's Directory of the Pulp, Paper and Allied Trades*. Miller Freeman, San Francisco

Casey J P (ed.) 1983 *Pulp and Paper: Chemistry and Chemical Technology*, 3rd edn. Wiley, New York

Gould F R 1976 Specialty papers. *Chem. Technol. Rev.* **78**, 246

Heimburger S 1998 Optimizing the TCF bleaching of a high kappa North American west coast softwood kraft pulp. *Proc. 1998 Pulping Conf.* TAPPI Press, Atlanta, GA, pt. 3

Hunter D 1978 *Papermaking: The History and Technique of an Ancient Craft*. Dover, New York

Lavigne J R 1993 *Pulp & Paper Dictionary*. Miller Freeman, San Francisco

Mosher R H, Davis D S (eds.) 1968 *Industrial and Specialty Papers*. Chemical Publishing, New York

Page D H 2000 The beating of chemical pulps—the action and its effects. In: Baker C F, Punton V W (eds.) *Trans. 9th Fundamental Research Symp. on Fundamentals of Papermaking*. Mechanical Engineering Publications, London, Vol. 1, pp. 1–38

Paulapuro H (ed.) 2000 Paper and board grades. In: *Papermaking Science and Technology*. TAPPI Press, Atlanta, GA, Vol. 18

Waterhouse J F, Barrett T D 1991 Aging characteristics of European handmade papers (1400–1800). *TAPPI J.* **74** (10), 207–12

J. F. Waterhouse
Peachtree City, Georgia, USA

Paper: Creep

The tendency of paper to conform to its surroundings is of great utility in the conversion of paper to an end-use product. Unfortunately, the same behavior can be detrimental to the end-use performance. An excellent example of this is the corrugated box. With the aid of steam and pressure, one can take a flat sheet and form the fluted medium that constitutes the middle layer of corrugated board. Then one can score and crease the combined board to produce a box, load it with goods, and store it in a warehouse at the bottom of a stack of similar boxes. Over time, the sides of the box will bulge, sometimes to the point where the stack collapses. All of these deformations are a consequence of the viscoelastic and plastic behavior of paper. The bulging of the box is a consequence of creep, which is defined as deformation accumulated over time and under constant load. The creep behavior of paper is of key importance to the paper and packaging industries.

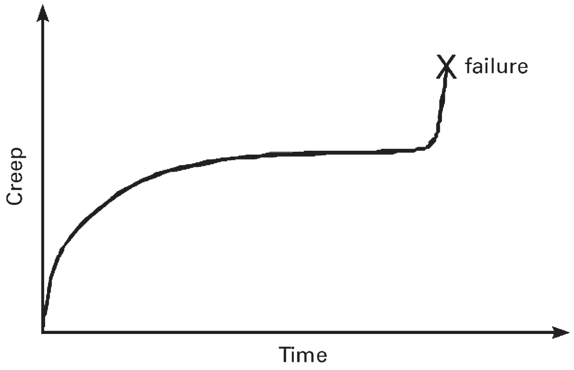


Figure 1
Typical creep response of paper.

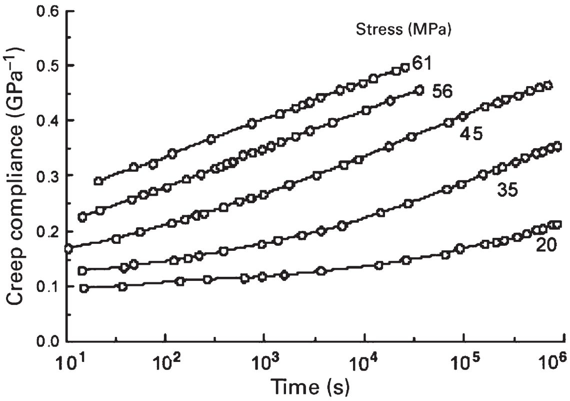


Figure 2
Family of creep curves used to form master creep curve (from data of Brezinski 1956)

1. Characteristic Creep Behavior of Paper

Paper, being a polymeric material, exhibits a creep response similar to that of many polymers. In addition, the creep behavior is affected by the fibers and chemicals used to make paper, environmental effects such as temperature and moisture, and the type and level of loading applied to the paper. The effect of moisture on creep is of special interest and the phenomenon known as "accelerated creep" is studied and discussed tirelessly in the paper industry. Brezinski (1956) completed a thorough scientific study of the creep behavior of paper, characterizing the creep behavior of paper in uniaxial tension as a function of applied load and moisture content of the sheet. All of his data exhibited the same characteristic behavior: when a sheet of paper is loaded, there is an immediate elastic response followed by creep, with a typical response as shown in Fig. 1. Note that the initial creep rate is rapid, but diminishes quickly with time. At some point, the creep rate can accelerate, leading to failure.

As shown in Fig. 2, Brezinski (1956) found that the creep compliance (creep strain divided by load) plotted against the logarithm of time for different load levels produced a family of curves that had log-linear regions of creep with equal slopes. Correctly shifting the curves in log time created a master creep curve that can be used to predict the creep strain at any load at any time. Brezinski (1956) also found a linear relationship between the shift in log time required to form a master creep curve and the applied load. The result is that paper exhibits a highly nonlinear relationship between creep rate and load.

A useful method for characterizing the creep behavior of paper is to form isochronous stress–strain curves, which are produced by plotting stress vs. strain at a specified time. The different points for the curve are obtained by varying the applied load and measuring the strain at the specified time. Isochronous curves are useful to discern the effect of varying a particular parameter. Such measurements were utilized in both tension and compression by Haraldsson *et al.* (1993). In Fig. 3, isochronous curves are shown for 1 and 10^4 s for both machine direction (MD) and cross-machine direction (CD) of a commercial linerboard subjected to edgewise compressive loading. Note that the curves are initially linear and become nonlinear with increasing strain.

If the load is removed after some creep has accumulated, partial creep recovery is observed. Paper also exhibits creep hardening, where the creep rate is reduced for subsequent creep loading of paper. It can be inferred from the work of Padanyi (1993) that much of the hardening effect can be eliminated by subjecting paper to a cycle of changing moisture content between creep loadings.

The creep response outlined above is typical of paper products. Specific papers will have slightly different behaviors, depending on their composition and testing conditions.

2. Compression vs. Tension

Haraldsson *et al.* (1993) found that compressive and tensile creep were the same at low loads, but as the load increased the creep was greater under compression. While a tensile creep test is very simple to perform, compression creep testing is difficult because of the low buckling resistance of a sheet of paper. This has led to several compressive creep tests, using lateral support fingers (Haraldsson *et al.* 1993), vacuum pressure to hold the specimen flat (Setterholm and Gunderson 1983), or testing of paper cylinders (Setterholm and Gunderson 1983). Although these tests produce the same basic behavior, the magnitude of results will most likely differ.

3. Effect of Various Parameters

Increasing moisture content of paper causes higher creep rates in both tension and compression (Brezinski 1956, Söremark *et al.* 1993). Brezinski (1956) also found that increasing the density of the sheet by wet pressing and beating reduced the rate of creep in paper. At the same density, increased beating of the pulp also reduced the creep.

In general, increased restraint during drying creates a sheet that exhibits less creep (Schulz 1961, Söremark *et al.* 1993). As shown in Fig. 3, machine-made papers typically exhibit less creep at a given load in MD than CD. This creep anisotropy is from the combined effects of drying restraint and fiber orientation (Söremark *et al.* 1993).

Byrd (1984) reported compressive creep rates of paperboard made from different fiber sources, including virgin, recycled, and high yield pulps that were very similar when tested under a load equal to 25% of the compressive strength of the board. Bobu *et al.* (1984) found that increasing calcium carbonate increased the creep rate of paper. Caulfield (1994) showed that use of cross-linking agents, such as formaldehyde, in the papermaking process greatly reduced creep rates.

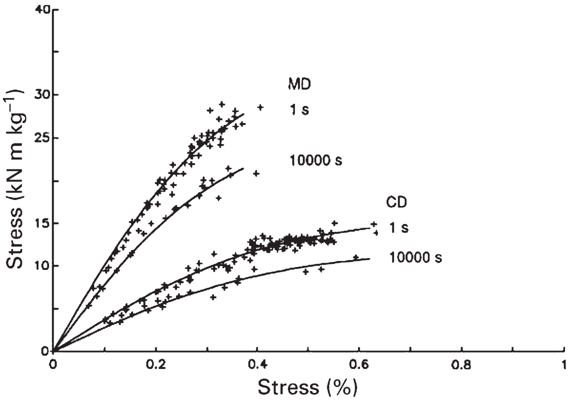


Figure 3
Example of isochronous creep curves for paper subjected to edgewise-compressive loading (from data of Haraldsson *et al.* 1993)

4. Accelerated Creep

Much of the creep research for paper has focused on the acceleration of creep by cyclic humidity. Since the first observations of Byrd (1972), it has been well established that a cyclic humidity environment is more severe for creep of paper than constant exposure to high relative humidity, which is illustrated in Fig. 4. This coupling of a mechanical action (creep) with a moisture effect (hygroexpansion) is termed the mechano-sorptive effect.

Accelerated or mechano-sorptive creep is probably the critical factor that limits the lifetime of a box. The majority of warehouses do not have controlled humidity environments, and the humidity cycles on a daily basis. This increases the rate at which the box creeps in compression and accelerates the bulging and failure of the boxes. Although the actual rate of accelerated creep varies as a function of the cycle time (Urbanik and Lee 1995), the total amount of creep appears to be relatively constant as a function of the number of cycles (Gunderson and Tobey 1990). Below some optimum cycle time, the accelerated creep rate decreases. This is especially true if the cycle

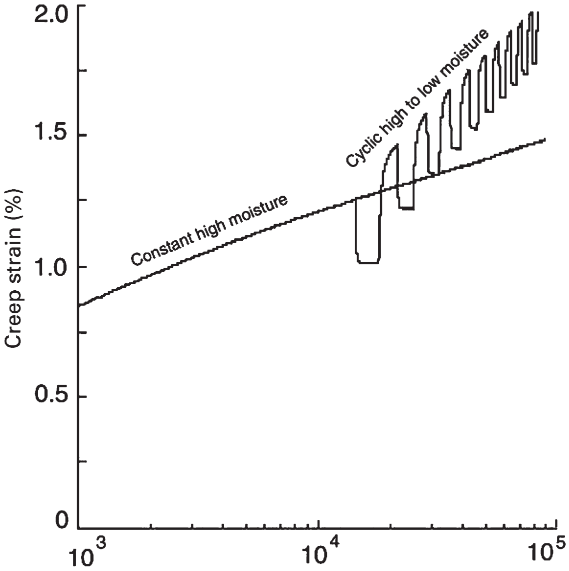


Figure 4
Creep at constant high moisture and accelerated creep of paper exposed to tensile loading (from data of Habeger and Coffin 2000).

time is less than the time needed for the paper to reach moisture equilibrium.

Researchers in the paper industry have focused on explaining the mechanisms of accelerated creep, relying on special phenomena such as dependence of creep rate on moisture gradients, or hydrogen bonds breaking and reforming. Söremark and Fellers (1993) have proposed that stress gradients and the formation of local dislocations explain accelerated creep. Habeger and Coffin (2000) have shown from a pure mechanics standpoint that accelerated creep is a natural consequence of the stress gradients created by changing moisture combined with the creep properties.

5. Summary

Creep behavior of paper is important to many converting operations and end-uses. Beyond creep, other rheological characterizations, including stress relaxation, dynamic mechanical analysis, and constant strain-rate loading and unloading tests, have been done for paper. The viscoelastic and plastic behavior of paper is quite difficult to fully characterize, because of the complexities arising from paper structure, temperature and moisture effects, differences in tension and compression, and the coupling of all these effects. There is still no adequate mathematical model that fully explains creep behavior. There is also not a full explanation of the mechanisms of creep and the role of fibers and bonded regions. One must still characterize creep behavior experimentally, including the mechano-sorptive effects. This is especially true when evaluating commercial paper for potential end-use products.

See also: Paper: Effects of Moisture and Temperature*; Wood: Dimensional Change from Moisture#; Paper: Strength and Stiffness*; Wood: Creep and Creep Rupture*

Bibliography

Brezinski J P 1956 The creep properties of paper. *Tappi J.* **39**, 116–28

Byrd V L 1972 Effect of relative humidity changes during creep on handsheet paper properties. *Tappi J.* **55**, 247–52

Byrd V L 1984 Edgewise compression creep of fiberboard components in a cyclic-relative-humidity environment. *Tappi J.* **67**, 86–90

Caulfield D F 1994 Ester cross-linking to improve wet-strength performance of paper using multifunctional carboxylic acids, butanetetracarboxylic and citric acid. *Tappi J.* **77**, 205–12

Gunderson D E, Tobey W E 1990 Tensile creep of paperboard—effect of humidity change rates. In: Caulfield D F, Passaretti J D Sobczynski S F (eds.) *Proc. Mater. Res. Soc. Symp. Materials Interactions Relevant to the Pulp, Paper, and Wood Industries.* Materials Research Society, Pittsburgh, PA, pp. 213–37

Habeger C C, Coffin D W 2000 The role of stress concentrations in accelerated creep and sorption-induced physical aging. *J. Pulp Paper Sci.* **26**, 145–57

Haraldsson T, Fellers C, Kolseth P 1993 The edgewise compression creep of paperboard—new principles of evaluation. In: *Trans. Fundamental Research Symp. Products of Papermaking*. Pira International, Oxford, pp. 601–37

Padanyi Z V 1993 Physical aging and glass transition effects on the mechanical properties of paper and board. In: *Trans. Fundamental Research Symp. Products of Papermaking*. Pira International, Oxford, UK, 521–45

Poppel E, Bobu E, Ciovica S 1984 Rheological behavior of papers produced with calcium carbonate fillers. *Zellstoff Papier*. **33**, 229–36

Schulz J H 1961 Effect of straining during drying on the mechanical and viscoelastic behavior of paper. *Tappi J.* **44**, 736–44

Setterholm V C, Gunderson D E 1983 Observations on load-deformation testing. In: Mark R E (ed.) *Handbook of Physical and Mechanical Testing of Paper and Paperboard.* Dekker, New York, Vol. 1, Chap. 4, pp. 115–43

Söremark C, Fellers C 1993 Mechano-sorptive creep and hygroexpansion of corrugated board in bending. *J. Pulp Paper Sci.* **19**, J19–25

Söremark C, Fellers C, Henriksson L 1993 Mechano-sorptive creep of paper—influence of drying restraint and fiber orientation. In: *Trans. Fundamental Research Symp. Products of Papermaking*. Pira International, Oxford, pp. 547–74

Urbanik T J, Lee K S 1995 Swept sine wave humidity cycle period effect on creep. *Wood and Fiber Sci.* **27**, 68–78

D. W. Coffin[a] and C. Fellers[b]
[a] *Institute of Paper Science and Technology, Atlanta, Georgia, USA*
[b] *Skogsindustrins Tekniska Forskningsinstitut, Stockholm, Sweden*

Paper: Effects of Moisture and Temperature

Paper physicists, in attempts to understand mechanical and rheological properties of paper, have used two apparently competing approaches—molecular and structural. The molecular approach assumes that the important mechanical and rheological properties of paper stem ultimately from the cohesive molecular forces that hold the material together. The structural approach views the important mechanical and rheological properties of paper as being controlled ultimately by fiber and network structure on a supermolecular level. In the structural approach the rheological behavior of paper arises from the bending, stretching, and shearing of fiber segments between bond sites as well as the deformation of fibers at the actual bond sites (Page 1969). This approach has included the many network models and deals with the structure of paper on a microscopic, rather than a molecular level.

These two views are not necessarily contradictory, but represent complementary views of paper as seen at different levels of magnification in its complicated hierarchy of structure. At an intermediate level of this hierarchy of structure is the macromolecular level, which provides an alternate view for explaining mechanical and rheological properties. Ultimately, although the mechanical properties must depend upon the intermolecular forces that hold any material together, there is a complicated interplay between forces and geometry or structure. Clearly the structure and morphology of paper provide the framework upon which molecules are arrayed and through which those intermolecular forces operate. Consequently, there are alternate explanations for the behavior of paper, dependent upon the level of the hierarchy of structure that is examined. However, in the important questions concerning the influence of temperature and moisture on paper properties, the defining level of structure is not arbitrary. How do temperature and moisture interact with paper? With what level of the hierarchy of the structure of paper does temperature or moisture interact? And is this interaction quantifiable? Clearly heating or cooling always involves a molecular process, and water interacts with cellulose or paper at a molecular level.

1. The Molecular Approach—Factoring Out Structure

The molecular approach attempts to describe the rheological behavior of paper in terms of the molecular forces. Macroscopic stress–strain behavior is assumed to arise from interatomic potentials and the straining of molecular bonds. A strong proponent of the molecular approach to explaining the rheology of paper has been Nissan (1977), who incorporated into his approach a number of simplifying assumptions to minimize mathematical formalism. The major simplification is that the only important interatomic potential is that associated with the hydrogen bond. Nissan's molecular approach emphasizing the importance of the hydrogen bond has been called the theory of hydrogen bond-dominated solids. (Nissan's many contributions to this theory span a period of almost 50 years.) A solid is generally considered hydrogen bond-dominated where stiffness and mechanical strength disappear when fully saturated with liquid water. Ice represents the pre-eminent hydrogen bond-dominated solid. Paper repulpability and recyclability offer evidence of being in this class of materials.

Nissan's hydrogen bond theory that deals primarily with the elastic modulus of paper (and to a lesser extent, the flow portion of the stress–strain curve) sees the hydrogen bond as the predominant average local molecular feature of the structure. It is distinct from structural models in that the elastic modulus of paper appears to be independent of its network structure. In its simplest form, assuming isotropy, Nissan's theory states that the elastic modulus of paper is proportional to the cube root of the number of effective hydrogen bonds per unit volume:

$$E = k(n)^{1/3} \tag{1}$$

where E is the modulus of elasticity, n is the number of effective hydrogen bonds per cubic volume, and k is the spring constant of the hydrogen bond. This expression states that the density of hydrogen bonds, along with the force constant of the hydrogen bond, determines the modulus of the material. This idea that the elastic modulus of a material depends only on intermolecular interactions is not unusual to polymer science. Polymer scientists have long understood that the elastic modulus of most polymers in the glassy state is determined primarily by the strength of intermolecular forces and not by the strength and geometry of the covalent bonds along polymer chains (Nielsen 1974).

Intermolecular forces are generally characterized by polymer scientists in terms of the material's *cohesive energy density* (CED), a measure of the energy needed to separate and remove completely one molecular unit from its neighbors. For small molecules, CED is approximately equal to the heat of vaporization at constant volume, but for polymers it must be evaluated by indirect methods. CED generally correlates with the elastic modulus of polymers. Tobolsky (1960) pointed out a simple proportionality between the bulk modulus of polymers and their cohesive energy densities and Holliday and White (1971) noted a similar proportionality between elastic modulus and CED of polymers. These observations that modulus is determined largely by intermolecular forces expressed in terms of cohesive energy density are very similar in essence to Nissan's suggestion that the elastic modulus of paper depends only on the density of hydrogen bonds, the type of intermolecular force that dominates in hydrogen bonded materials. Two aspects of paper that the molecular or hydrogen bond approach can help explain in both qualitative and quantitative terms are the effects on elastic modulus that are caused by changes in either temperature or moisture.

2. Effect of Temperature on Modulus

Nissan (1977) has provided in some detail a derivation of the temperature dependence of elastic modulus (see also Dodson and Herdman (1982) for a concise summary.) Using statistical mechanics and the well founded Lippencott-Schroeder potential for the hydrogen bond, Nissan and Batten predicted a temperature dependence for the elastic modulus of cellulose to be approximately equal to

$$\mathrm{d}(\ln E)/\mathrm{d}T \sim -2.4 \times 10^{-3}\ {}^{\circ}\mathrm{C}^{-1} \tag{2}$$

This relationship was rigorously derived without arbitrary parameters (Batten and Nissan 1987) and means that the elastic modulus should decrease about 0.24% for every 1 °C rise in temperature. Nissan (1977) compared the predicted value with measured values for a variety of cellulosics and found the theory reasonably predictive. Values for $d(\ln E)/dT$ obtained from the literature ranged from -1.8×10^{-3} to $-6.3 \times 10^{-3}\,°C^{-1}$ for various cellulosics including paper. Two additional sets of experimental data on paper also give values for the temperature dependence of modulus that support the theoretically derived value. Salmén and Back (1977) report that between −25 °C and 175 °C the elastic modulus of paper decreases at a rate of about 0.3% per degree. From data obtained in torsional experiments, de Ruvo *et al.* (1976) report the relative change in rigidity of paper with temperature to be about $-2 \times 10^{-3}\,°C^{-1}$. No network theory of paper has been capable of predicting a temperature dependence of modulus.

3. Effect of Moisture Content on Elastic Modulus of Paper

The reduction of the elastic modulus with moisture is another property of paper that is well described by Nissan's theory of hydrogen bonded solids. This reduction in elastic modulus (E) is viewed as resulting simply from the reduction in the number density of effective hydrogen bonds by the disrupting, competitive action of water. Nissan concludes that the elastic modulus should show an exponential decline with moisture content, w, in two distinct regimes.

$$\text{Regime I for}(0 < w < w_{\text{crit}})\ d(\ln E)/dw = -1 \qquad (3)$$

$$\text{Regime II for } w_{\text{crit}} < w\ d(\ln E)/dw = -(\text{C.I.}) \qquad (4)$$

In the region of low moisture contents, below monolayer coverage, w_{crit} (about 5% moisture content), the value of the slope when plotting in (modulus) versus moisture content is expected to be about minus unity. At moisture contents above w_{crit}, the negative slope is greater (equal to ~5 or 6), and is designated by the symbol C.I., signifying *Cooperative Index*. Nissan views the reduction of E with increasing moisture as resulting simply from a reduction of the number of effective network hydrogen bonds by the disrupting, competitive action of water.

In the region of low moisture contents, below about 5% moisture, the value of unity is predicted. At moisture contents above ~5% (and to saturation), the value of C.I. is considerably larger; predicted by Nissan to be about 6.7. This value was obtained by combining the ideas of hydrogen bond-dominated solids with the cluster integral concept of water in order to help explain how one molecule of water added to a cellulosic network seems to disrupt several effective network bonds; i.e., a cooperative bond-breaking effect. A value of 6.7 for C.I. indicates that the decrease in elastic modulus is expected to be about 7% for every 1% increase in moisture content. Data at low moisture contents are sparse, but for moisture contents above 5%, Nissan (1977) has compared the experimental value of C.I. evaluated from results reported in the literature and, excluding values obtained from sonic measurements, finds values ranging from 5.7 to 8.2. Sonic and other dynamic methods of measurement (Caulfield and Weatherwax 1978, Htun *et al.* 1985) have shown exponential decrease of elastic modulus with increasing moisture content, but give smaller values for C.I., in the range of ~3–5.

3.1 Moisture Effect on Ultrasonic versus Static Modulus Measurements

For many years there have been these unresolved differences between measurements of elastic modulus. Typically the values for modulus obtained ultrasonically, E_u, have exceeded those measured on mechanical-type machines, E_s (the subscript s signifying static as opposed to the ultrasonic {u} method). At low moisture content, the elastic modulus measured sonically, E_u, exceeded the static modulus E_s, by about 10%. But this discrepancy between measured E_u and E_s appeared to worsen as the moisture content of the paper increased. Nordvall (1992) has confirmed these differences and has provided data comparing elastic modulus measured by both methods measured on 24 different papers in both machine and cross-machine directions as a function of moisture content. These results clearly quantify the nature of the differences, and have provided important data with which to test some aspects of the internal consistency of one of the predictions of the hydrogen bond theory.

One starts with the basic assumption that the elastic modulus depends on the cube root of the effective number density of hydrogen bonds (Eqn. (1)) and that the effective number density of hydrogen bonds is different. Depending on the method of measurement; that $(n)_u > (n)_s$, and that the observed moisture dependence of the measured modulus is given by the relationship shown in Eqn. (4), then it can be shown (Zauscher *et al.* 1996, 1997) that

$$(E_0)^3_u/(E_0)^3_s = \{(\text{C.I.})_s + 1\}/\{(\text{C.I.})_u + 1\} > 1 \qquad (5)$$

The cube of the modulus ratio (ultrasonic/static) at zero moisture content should be predictable from the ratio obtained from the slopes of plots of ln$E$ versus moisture content in Regime II. Although a ratio larger than unity arises from the initial assumptions, the prediction that the cube of the modulus ratio should be derivable from the moisture dependence of the modulus can only arise if the concepts of the hydrogen bond theory are valid.

The two ratios were obtained from the experimental measurements in both machine and cross-machine directions, and also for equivalent isotropic papers. For both ratios the values are greater than unity, closer to 1.3, but they are not statistically different. This equality of the ratios indicates an internal consistency with the hydrogen bond explanation of the phenomenon. The experimental results also indicate for typical papers as those examined, that in the dry state, sonically measured elastic modulus exceeds the statically determined modulus by about 10%, but that the difference becomes more pronounced, i.e., $>30\%$, as moisture content increases. These differences arise from the existence of two different forms of energy storage and propagation in paper dependent on the method and time scale of modulus measurement.

3.2 Nonlinear and Viscoelastic Effects on Strength of Paper

When paper is deformed by external forces, part of the work done is dissipated as heat, but paper can also store and dissipate mechanical energy. If a solid is subjected to a sinusoidally oscillating stress, and the strain is intermediate between exactly in phase and 90° out of phase with the stress, then its behavior is called viscoelastic. Paper behavior is viscoelastic, which is evident in other rheological properties and stress–strain behavior. Paper exhibits nonlinear load–elongation curves. Paper creeps when subjected to a constant load; deformation of a paper strip is not instantaneous, but elongates with time. Analogously, paper also exhibits stress-relaxation. If paper is stressed and held at a fixed elongation, the internal stresses relax with time. Paper also exhibits marked rate of loading (ROL) and duration of load (DOL) effects (Caulfield 1985). The measured strength of a piece of paper increases with the rate of stressing. The failure strength in a linear-ramp ROL experiment increases with the logarithm of the rate of stressing. In DOL behavior, paper deadloaded to a level below its normal breaking stress will nevertheless fail if the stress is maintained for sufficient time. The logarithm of the time to failure under a constant deadload stress increases with decreasing stress levels.

The presence of moisture in paper accentuates all viscoelastic characteristics. In addition, moist paper and paper in changing humid environments exhibit a host of additional viscoelastic phenomena generally classified as mechanosorptive effects. Grossman (1976) has described a long list of these effects as they are manifested in wood, and many presentations in one of a series of symposia (Fellers and Laufenberg 1994) have been devoted to topics of mechanosorptive effects in paper.

3.3 Viscoelastic Properties

The elastic properties of the hydrogen bond-dominated solid are described as dependent on an average number density of effective hydrogen bonds. Little is said about its time dependence and how this average is obtained. In order to make explicit the viscoelastic properties of a hydrogen bond-dominated solid, the labile nature and time dependence of the hydrogen bond distributions must be considered. Molecular models show that the cellulose molecule is completely covered with hydroxyl groups. Within the fiber, the stiff cellulose molecules are connected in long chains by strong chemical covalent bonds. But these chains are held side-by-side by hydrogen bonds that are much weaker than the chemical bonds along the chains. Within the cellulose structure there is little chain entanglement and there are no chemical covalent bonds between fibers, also held together by only weak hydrogen bonds. The hydrogen bond is so much weaker than the covalent bond that its energy is of the same order of magnitude as the thermal kinetic energy associated with molecules at ordinary room temperature:

Covalent bond energy $340\,\mathrm{kJ\,mol^{-1}\,K}$
Hydrogen bond energy $\sim 19\,\mathrm{kJ\,mol^{-1}\,K}$
Thermal kinetic energy at room temperature $3.8\,\mathrm{kJ\,mol^{-1}\,K}$

Only 1% of molecules at ordinary temperatures (estimated by Boltzmann statistics) have thermal energies greater than the energy associated with the hydrogen bond. This can only mean that hydrogen bonds are breaking and forming very quickly at room temperature, just as they are presumed to do in water. Therefore, paper is held together not by a locked array of hydrogen bonds, but rather a flickering array of bonds. At any one instant, a fraction of the hydrogen bonds is open. This flickering array is effective in providing the mechanical integrity and high modulus of paper since the array is densely spread over a large area, and because cellulose chains are stiff. Even if a number of bonds open cooperatively while a dry fiber assembly is stressed, the resultant strains can be small because the hydrogen bond array tends to reform quickly with the same atoms paired.

When paper is exposed to moisture or high humidity, cellulose absorbs water between fibers and within the cell wall (Chatterjee and Ramarao 1997). The water molecules compete for the hydrogen bonding sites, disrupting bonds between cellulose surfaces. The swelling action of water separates the cellulose chains from one another (and at the same time plasticizes them, making them more flexible by allowing more free volume.) Now if a stress is applied when water is present, the hydrogen bond array still flickers open and closed. But on closing, the array need not reform with its original partners paired. The details of how the flickering array responds must clearly depend upon the time duration over which the stress is applied. In this respect, the number density of effective hydrogen bonds has an inherent time dependence. Two instances in which this time dependence is revealed are (i) the differences in elastic modulus when measured by ultrasonic or static methods (mentioned above), and (ii) creep and stress relaxation.

3.4 Bond-breakage Theories, Stress Relaxation and Creep of Paper

When a sheet of paper is subjected to a constant stress, it will deform and the strain will increase with time. Likewise, when a sheet of paper is strained and held at constant length, the initial stress relaxes with time. Rance (1950) was the first to suggest that the irreversible flow characteristics of the viscoelastic behavior arise from a continuous series of internal ruptures in the sheet, and ultimately lead to failure. However, the irreversible breakage of bonded area between fibers has been reported to play only a modest role, and then only in the final stages of the stress–strain curve near failure. Steenberg's (1947) idea of viscous elements was introduced to the bond-breaking theory when the bonds are weak and have the potential to reform after breaking. This concept led to the hydrogen bond explanation of elastic and viscoelastic behavior. Although hydrogen bonding alone may be the predominant structural entity during the early stages of deformation, and thereby easily related to tensile modulus, it is difficult to relate hydrogen bonding with rheological behavior further out on the stress–strain curve, where other features of fiber and network structure contribute. However, Nissan and co-workers have made attempts to relate the plastic or nonlinear region of the stress–strain curve to the form of the hydrogen bond potential (Nissan *et al.* 1985).

Stress relaxation, (or equivalently creep) has been described by Nissan (1977) in terms of the kinetics of hydrogen bond breaking and reforming. He viewed stressed hydrogen bonds as being less stable than unstressed, so that stressed bonds flicker open and re-close in different unstressed configurations. And because this stress relaxation is viewed in terms of chemical bond rearrangements, classical chemical kinetics can be applied. If the relaxation is assumed to result from the cooperative process in which bonds break simultaneously, Nissan suggested a basic kinetic equation (Nissan 1967):

$$-\mathrm{d}(n/n_0)/\mathrm{d}t = k(n/n_0)^a \qquad (6)$$

The left side of the equation represents the rate at which the fraction of the initially stressed bonds relax into an unstressed state; k represents the 'chemical' rate constant; and a, the order of the reaction, is the C.I. of the bonds breaking simultaneously. Analysis of the experimental data led Sternstein (1961) and Nissan and Sternstein (1962, 1964) to conclude that there was more likely a spectrum of C.I. involved. For example, stress relaxation data for cellulose at 75% RH exhibits C.I. maxima at about −70 °C and −150 °C, i.e., close to temperatures at which instrumental methods also indicate relaxation phenomena. This interpretation of stress relaxation, however, is based on the axiom that viscoelastic properties can be interpreted in terms of the hydrogen bond. From this viewpoint, viscous effects are the result of the time-dependent change in the number of effective hydrogen bonds, and not primarily due to the flow characteristics of a viscous element.

4. Factoring in Structural Considerations

4.1 Models for Viscoelasticity

The study of the viscoelastic behavior of materials may be approached on many levels. For much of its history, one approach to viscoelasticity has been in terms of models. Springs and dashpots coupled as Maxwell bodies and/or Voigt (Kelvin) bodies and, arranged in a multitude of arrays, have been used as helpful analogs for different viscoelastic materials (Alfrey 1948) including paper. Dashpots containing non-Newtonian fluids have often been found necessary to more accurately represent the behavior of the model to a real viscoelastic material (Andersson *et al.* 1949). In all cases, however, the use of spring/dashpot models, combining elastic and viscous elements, is essentially an empirical curve-fitting technique, and far removed from actual molecular behavior.

Aside from the thermodynamic approach to rheology discussed later, viscoelasticity may be described from a fundamental phenomenological approach using only mathematics and a continuum theory. To some purists, this is the only sound way that viscoelasticity can be understood. However justified that viewpoint may appear, the fact remains that matter (and paper certainly) is not a continuum but consists of molecules, and the differences in properties of different materials must be due essentially to the differences in their molecular structures and behavior.

4.2 Macromolecular Behavior

One molecular approach to describe rheological behavior of paper in terms of molecular structure applies concepts that have been successful in the field of macromolecules; glass/rubber transitions, free volume, and plasticization. The methods have been adopted from polymer physicists, where dynamic mechanical behavior, dielectric behavior, ultrasonic response behavior, dilatometry, and DSC properties have been used. In the dynamic response methods, behavior is monitored as frequency, temperature, and/or moisture are varied. To the polymer physicist dealing with synthetic polymers, damping peaks in many macromolecules can be identified as belonging to certain specific atomic groups. For synthetic polymers, these results are presently the best examples of experimentally observed connections between rheological behavior and molecular dynamics.

As strains become larger, however, the role of other elements of structure necessarily increases. The function of cellulose chains can be looked upon as providing a framework, a geometric backbone, upon which the hydroxyl groups are arrayed. The morphology and structure of fibers determines how this framework is distributed through space and, thereby, how the hydrogen bonds are distributed. As strains increase, distortions arise not only from increases in interatomic separations, but also from movement of larger elements of the framework structure. Consideration of larger elements of structure also permits the treatment of anisotropy, whereas the molecular approach fits easily only with an idealized case of uniform distribution in three dimensions.

One way to factor in the role of polymeric and fiber structure has been through previously mentioned methods in polymer research, permitting combining molecular chain mobility and the effects of both temperature and moisture on mechanical behavior of paper.

To the paper physicist, the problem of relating rheological behavior to molecular structural dynamics has proven to be more difficult. Cellulose appears to be more complicated than common synthetic polymers; experimental results concerning various measured transitions are incomplete and, in certain respects, not reproducible (Klason *et al.* 1976). In cellulosics, there is general agreement that two relaxations are observed below room temperature at low frequencies (~ 1 Hz) at about 250 and 180 K. These have been suggested as arising from conformational changes involving methylol groups (Bradley and Carr 1976). Relaxations reported near 373 K are probably due to evaporation of water; and the weakly developed relaxations near room temperature have been attributed to transitions of adsorbed water or to other conformational transitions that were first suggested by Wahba (1965).

It is not until considerably higher temperatures are reached that transitions appear like true glass transition. Salmèn and Back (1977, 1980) in quantifying the effect of temperature on elastic modulus, found a marked decrease in modulus starting at about 200 °C that resembled a glass transition. They found that for dry cellulose the glass transition is as high as 230 °C, while for hemicelluloses it lies between 165 °C and 225 °C. They describe the effect of water as plasticizer and interpret its effect on the mechanical properties of cellulose as a lowering at the effective glass transition of the polymer. Qualitatively this must be the case, but there is no indication as to what molecular chain segments might become mobile, and at what moisture contents. Salmèn and Back admit that an interrelation must exist between apparent glass transition and a pronounced increase in hydrogen bond mobility. Goring (1966) has suggested a description in which water does not plasticize cellulose at a molecular level, but rather at a microfibrillar level.

Kubàt and co-workers (1969) and de Ruvo and co-workers (1976) have also related elastic modulus to volume effects that result from either thermal expansion or hygroexpansivity. Paper coefficient of thermal expansivity is of the order of $1 \times 10^{-5}\,°\mathrm{C}^{-1}$, and is approximately three times greater in CD than MD. Kubàt and co-workers (1969) found that the product of the elastic modulus and the thermal expansion coefficient is constant and independent of direction in which the measurements are made. De Ruvo and co-workers (1976) verified this observation and also found a similar approximate relationship for swelling. The product of hygroexpansivity and elastic modulus was also almost constant, independent of the direction of measurement. The swelling relationship appeared somewhat improved, however, if the geometric mean of the dimensional swelling was combined with the geometric mean of the elastic modulus in the MD and CD directions. From the similarity of behavior of paper sheets and individual fibers, de Ruvo *et al.* (1976) concluded that the network structure *per se* has little role in changes caused by temperature or moisture on elastic properties. Instead, response to changes in temperature and moisture are governed by properties and structures intrinsic to fibers and not a result of their network consolidation.

The addition of plasticizing water is qualitatively equivalent to an increase in temperature in that it provides an increase in the free volume available. Expansional strains, caused either thermally or by swelling, are inversely related to elastic modulus.

As the distance of separation between interacting groups is increased, either by thermal expansion or by swelling, the mean attractive force diminishes. This reduced attraction can be viewed as a type of bond weakening, and, in the limit, as a type of bond breaking. In this way, the concepts of glass transitions, plasticizers, and free volume afford a means of rationally connecting the mechanical properties of an anisotropic fiber structure to one whose elastic properties may also be described as hydrogen bond-dominated.

5. *Ultimate Strength Properties*

Paper strength properties as evaluated from tests to failure or rupture are more difficult to relate to the structural features of paper. Whereas elastic modulus reflects an intrinsic, average structural feature of a sheet, ultimate strength properties may be dominated by other less general features such as flaws and defects. Nevertheless, strength properties show many of the same correlations with temperature and moisture as the elastic modulus. (The survey by Crook and Bennett (1962) remains the best overall source of general data on the effects of humidity and temperature on paper's ultimate strength properties.)

The effects of temperature on ultimate strength properties tend to be less important than the effects of moisture and are often masked by the small changes in RH or moisture that accompany changes in temperature. Wink (1961) has presented a graphical representation of the average effect of temperature on strength of paper while keeping moisture content of the paper constant (by adjusting RH). Over a range of about 30 °C, he shows that tensile strength drops about 20%, or at a rate comparable to the loss in tensile modulus. Stretch, on the other hand, shows an increase of about 30%. Tensile energy absorption, being effectively the integration of tensile stress over strain up to failure, shows only modest changes ($<10\%$) in the same temperature range.

The generally held view is that paper strength varies with temperature and moisture content by affecting two basic factors. Page (1969) has developed a network theory of tensile strength that identifies and tries to separate the effects of these two factors, fiber-to-fiber bonding and individual fiber strength. Fiber bonding and fiber strength do not react similarly to moisture. Fiber bonding weakens as moisture is absorbed, while individual fibers become softer, more pliable, and stronger (Crook and Bennett 1962). Physical properties that are measures of strength tend to have maxima at low or moderate relative humidity (RH). Properties for which flexibility is of importance will generally increase in value as humidity rises. Tensile strength, for example, tends to have a maximum at RH between 30% and 50%. Above 50% RH, the tensile strength tends to show a moderate decrease to about 65% RH, where it drops sharply with further increase in RH. Extensibility or strain to failure, conversely, tends to increase continuously over the full range of increasing RH. Strain to failure may increase over four fold from 20% to 90% RH, an increase among the highest of any mechanical property. Burst strength is a complex property that is influenced by both tensile strength and extensibility; and a change in moisture affects the tensile strength and extensibility in opposite directions. Thus, of strength tests, burst shows the least effect of increasing humidity. A typical curve of burst value vs RH shows little curvature, with only a slight maximum between about 40 and 70% RH. Tear resistance generally shows a continuous increase of between 0.5 to 1% per $+1\%$ RH to about 80% RH. At very high humidities, tear resistance also decreases as very weakened fiber–fiber bonding allows fibers to slip apart more easily. Fold endurance is a measure of flexibility and tends to show an increase of between 1 and 5% per $+1\%$ RH to above 80% RH. At very high humidities the small tensile load applied to the fold specimen probably becomes limiting and fold endurance also drops off near saturation.

6. *Time-dependent Effects*

The most challenging unresolved problems in the rheology of paper are connected to time effects. Paper exhibits a wide range of time-related phenomena, and the introduction of time into the already complicated interrelationships of temperature, moisture, and paper structure increases the difficulty. Kolseth and de Ruvo (1983) have reviewed the time–temperature–moisture-dependent properties of paper, which is the most current and comprehensive in the field. There are three main areas of study in which time effects are significant: (i) creep and stress relaxation, (ii) duration lifetime/rate-of-loading phenomena, and (iii) dynamic mechanical measurements.

6.1 Duration-of-load/Rate-of-loading Phenomena

Paper exhibits two related time-dependent rupture phenomena that influence its load-bearing capacity. If paper is stressed to only a fraction of its short-term breaking stress and maintained at that stress over a long enough time, the paper will fail. It has been found empirically that for paper under constant stress, the logarithm of the time to failure varies inversely as the stress level (Guthrie and Fulmer 1969). A related time-dependent phenomenon is the rate-of-loading effect. The breaking stress of paper increases as one increases the rate at which it is stressed. If the applied stress is increased linearly with time, the breaking stress increases with the logarithm of the rate of stressing (Rance 1953). For this reason meaningful strength testing always requires a standardized rate of stressing (or straining).

Duration lifetime studies of materials can be aided by the use of a so-called *semi-empirical* equation that has been credited to Zhurkov (1965).

$$t = t_0 \exp[(l/kT)\ (U - \gamma\sigma)] \tag{7}$$

where t is the time to failure, T the absolute temperature, k Boltzmann's constant, σ stress, U activation energy, and t_0 and γ are constants.

A form of Eqn. (7) was derived and used years earlier by Coleman (1956) to describe the duration-of-load and rate-of-loading phenomena exhibited by nylon. In Coleman's analysis, the equations describing the duration-of-load behavior and rate-of-loading behavior arise from the integration of an equation he derived for the rate-of-creep deformation. Coleman's fundamental equation, based on Eyring's rate theory, involves an explicit temperature dependence (for details see Coleman's classic paper, 1956). At fixed stress, Coleman's equation indicates that increased temperature will decrease a loaded specimen's lifetime, and this behavior has been confirmed for viscose yarn as well as nylon.

De Ruvo and Lundberg (1972) have reported the moisture dependence of duration lifetime for chipboard at three RH. At a constant stress level, paper at a low RH will support a stress for a longer time before failing than at a high RH. Alternatively, paper at a high RH must be loaded to a smaller stress level in order to equal the duration lifetime of the same paper at low RH. One area of intense research involves the phenomenon of accelerated creep in cyclic humidity conditions (Fellers and Laufenberg 1994, de Ruvo and Lundberg 1972, Byrd 1972a, 1972b, Gunderson 1981, Back *et al.* 1983). Whereas paper may exhibit a creep rate at 90% RH that is several times the creep rate at 35% RH, if the same loaded paper sample is cycled between 35% and 90% RH, its creep rate can be over 10 times the creep rate shown at the constant 90% RH (Sternstein 1961).

In a phenomenon related to the case of accelerated creep in cyclic humidity, Byrd (1972a, 1972b) has reported the shortened lifetimes in both tension and compression of paper loaded in cyclic humidities. Similar findings were reported also by de Ruvo and Lundberg (1972). A better understanding of this phenomenon of accelerated creep-rupture is crucial in the design of packaging and in the structural performance of paper in long-term exposures to cyclic-humidity environments. The basic features of creep and stress relaxation of paper in tension and compression are similar. However, the effects appear to be more serious in the case of compression.

This phenomenon of creep-rupture in cyclic humidity probably arises because of the interference or coupling of two thermodynamically irreversible processes: sorption hysteresis and stress–strain hysteresis. The thermodynamic fundamentals of the coupling of deformation and sorption, the so-called *mechanosorptive* effect, was provided by Barkas (1953); but a complete description of the irreversible thermodynamics involved by the methods of either Prigogine and DeFay (1954) or Onsager (1931) is lacking. In mechanistic terms, the phenomenon is describable in terms consistent with the theory of hydrogen bond-dominated solids (Caulfield 1980). During each adsorption leg of a humidity cycle, a 'spreading pressure' pries apart two cellulose lamellas and intercalates a layer of water molecules, reducing the number of effective load-bearing elements. The subsequent reduction in humidity fails to reverse completely this process, because the restraining modulus is lower at the higher moisture content, and also stressed bonds have had time to re-form into stress-relaxed configurations. Failure can then be triggered, as in Coleman's creep–rupture theory, when the strain deformation exceeds a critical strain. A better creep–rupture explanation will probably involve consideration of the accumulation of nonrecoverable work, and the formulation of fatigue lifetime in terms of energy dissipation rather than as a result of a critical strain (Caulfield 1994).

6.2 Dynamic Mechanical Behavior (Fatigue) of Paper

When paper is subjected to a periodic stress, its strain response exhibits a time-dependent behavior that arises from its viscoelastic nature. If a simple harmonic stress is applied to paper, its strain response is also periodic but lags behind the stress. This time lag can be measured in terms of a phase angle, δ, whose tangent is equal to the ratio of loss modulus to storage modulus. Tan δ is a measure of the energy lost due to the internal friction generated in the sample during its cyclic deformation. When examined as a function of frequency of oscillation, the internal friction or tan δ exhibits a maximum wherever the elastic modulus undergoes a rapid change. In polymers, the frequencies at which energy losses are maximized have been interpreted as caused by transitions from glassy to rubbery behavior. The wood polymers, lignin and hemicelluloses, have been recognized as essentially thermoplastic polymers whose glass–rubber transitions are both temperature- and moisture-dependent (Goring 1966). For example, whereas isolated dry lignin softens at a temperature near 200 °C, in water its transition is lowered to about 80–90 °C. Hemicelluloses are even more moisture sensitive; softening at about 55 °C at a moisture content of only 23%. The work of Kubàt (1969), de Ruvo *et al.* (1976), and Htun *et al.* (1985) has verified that both increased temperature and moisture dampen the dynamic response of paper to cyclic stress.

An area of intense research interest has been the effect of changing moisture content on the dynamic loss property of paper (Htun *et al.* 1985). Kubàt and Lindbergson (1965) first reported that a sudden change in humidity causes a transient increase in mechanical damping and suggested that the damping was associated with the moisture diffusion process taking place during equilibration. An analysis by Salmèn and Back (1977) has discounted the importance of moisture gradients *per se* but emphasizes the temporal change of gradients with time. Htun *et al.* (1985) have studied the dynamic mechanical behavior of paper during drying and attribute transient damping as well as increased modulus to the loss of mobility of amorphous components of cell wall polymers. Understanding the dynamic behavior of paper in transient-humidity environments will certainly aid our understanding of how creep and failure are accelerated by cyclic humidity. Whereas dynamic behavior may be explained phenomenologically in terms of chain motions and second order transitions, there is good reason to believe that this time-dependent phenomenon, too, is ultimately governed on the molecular level by the making and breaking of hydrogen bonds.

7. *Thermodynamics Approach*

Finally, a thermodynamic approach formulated by Meixner (1960) shows that the introduction of an additional time-dependent internal variable is both necessary and sufficient to provide a complete description of viscoelastic behavior. A correspondence between Meixner's model-less method and the suggested time dependence of the number density of effective hydrogen bonds has been suggested (Caulfield and Zauscher 1997).

For a general presentation of viscoelasticity of paper in terms of irreversible thermodynamics, Meixner (1960) has done this rigorously, and Kestin (1978) has put Meixner's ideas into comprehensible terms. This treatment not only has a general validity, but also results in a form that seems directly applicable to the earlier analysis of the differences between ultrasonic and 'static' measures of elastic modulus.

Kestin (1978) starts with the argument (when considering creep and relaxation) that the usual number of variables needed to define the state of a material is insufficient. A stressed material maintained at a fixed elongation and temperature, for instance, does not exhibit a constant stress if observed over long times, but relaxes. At least one other degree of freedom is necessary to define its state. He calls that other variable, ξ; an internal degree of freedom or parameter of state. Prigogine and DeFay (1954) called such an additional variable a "degree of advancement" especially when the change involves a chemical change or chemical reaction in the material. Although we do not know what change the variable represents, it must exist from Meixner's thermodynamic analysis. For our purposes, it represents the degree of advancement of the 'chemical' change involving transitions of bonds from effective (load-bearing) to noneffective (nonload-bearing). However, we never know from thermodynamics what ξ might represent, because thermodynamics is independent of any putative model. Kestin establishes the case for viscoelasticity by defining the Helmholtz function, F, as a function of strain, ε, and that additional variable, ξ at constant temperature as a Taylor series expansion about equilibrium conditions:

$$F(\varepsilon,\ \xi) = \tfrac{1}{2} E_\infty \varepsilon^2 + b\varepsilon\xi + \tfrac{1}{2} c\zeta^2 \qquad (8)$$

where E_∞ is the elastic modulus at infinite strain rate (because ε does not have time to change) and b and c are constants characteristic of the material. Because

$$\mathrm{d}F = \sigma \mathrm{d}\varepsilon - A\mathrm{d}\zeta \qquad (9)$$

then stress,

$$\sigma = (\partial F/\partial\varepsilon)_\xi = E_\infty \varepsilon + b\xi \qquad (10)$$

and the associated affinity,

$$-A = (\partial F/\partial\xi)\hat{\varepsilon} = b\varepsilon + c\xi \qquad (11)$$

One of the immediate results

$$\sigma = (E_\infty - b^2/c)\varepsilon = E\varepsilon \qquad (12)$$

where E is the elastic modulus at a lower rate of straining, and necessarily $E_\infty > E$. In this form the similarity of the relation between ultrasonic modulus and static modulus is apparent (compare Eqn. (5) where E_u is analogous to E_∞). Further research is needed to confirm these similarities and quantify the magnitudes of the material properties, b and c. Kestin also shows, using Meixner's arguments, that through the introduction of just this additional variable, ξ, a complete, consistent description of the phenomena of stress relaxation and creep is achieved.

The model-less thermodynamic approach of Meixner shows that a description of viscoelasticity arises from an additional internal variable that has a time-dependent character. By allowing a time dependence to the effective number density of hydrogen bonds, the hydrogen bond theory that accounts well for the elastic properties of paper can be extended to explain the viscoelastic properties of paper as well. Alternatively, to the polymer chemist, the explanation of relaxation phenomena involves molecular motions that result from the increased free volume that arises either from thermal expansion or the introduction of plasticizer molecules. For cellulose and paper, at usual temperatures and moisture contents, it is clear that significant hydrogen bond disruption must precede any free-volume increases associated with glass-like transitions. In papers with high noncellulosic components, at high moisture contents, the assumption that mechanical properties are controlled only by hydrogen bonds becomes less valid. But even in those cases, to the extent that increased molecular motion requires an increase in free volume, the basic concepts that apply to true 'hydrogen bond-dominated' solids must be valid. An advantage to the hydrogen bond explanation is its quantitative aspect of relating mechanical properties to intermolecular attractions, the ultimate source of all mechanical properties.

8. *Concluding Remarks*

A number of materials science approaches have been employed to understand the mechanical and rheological behavior of paper. These have included both network theories and molecular theories and both have shown some measure of success. However for understanding the important time–temperature–moisture relationships that affect elastic and viscoelastic behavior, the molecular approaches have been more successful because they examine the role of structure of paper at the appropriate level of interaction, where hydrogen bonds dominate interactions. A molecular approach called the theory of hydrogen bond-dominated solids has succeeded in describing in both qualitative and quantitative terms the effects of temperature and moisture on paper's elastic properties. Paper's viscoelastic properties are also adequately described qualitatively by this hydrogen bond approach. The importance of paper's macromolecular structure is emphasized by the materials science concepts of glass transitions, free volume and plasticization that help explain the roles of temperature and moisture in affecting paper's properties. The ultimate strength of paper, its inelastic deformation, creep, and rupture properties are greatly dependent on morphological and network characteristics. But how these properties are affected by temperature and moisture is through paper's macromolecular nature and largely through processes that involve the making and breaking of hydrogen bonds. Finally thermodynamic (model-less) approaches to rheology that make no materials science assumptions about the nature or structure of the material have displaced earlier empirical models, such as springs and dashpots.

Acknowledgments

The research community was saddened by the recent death of Alfred Nissan on 24 March 2000.

See also: Paper, Wetting and Penetration of Liquids into$^{\#}$; Paper: Nondestructive Evaluation$^{\#}$; Paper: Strength and Stiffness*; Wood: Moisture Content, Hydroscopicity, and Sorption*

Bibliography

Alfrey T 1948 *Mechanical Behavior of High Polymers*. High Polymer Series No. 6, Interscience

Andersson O, Ivarsson B, Nissan A H, Steenberg B 1949 The practical significance of paper rheology. *Proc. Techn. Sect. Paper Makers' Assoc. Gt. Britain* **30**, 43 New York

Back E L, Salmèn L, Richardson G 1983 Transient effects of moisture sorption on the strength properties of paper and wood-based materials. *Svensk Papperstidn*. **86**, R61–71

Barkas W W 1953 The mechanical properties of wood and their relation to moisture. In: Meredith R (ed.) *Mechanical Properties of Wood and Paper*. North Holland, New York

Batten G, Nissan A H 1987 Comments on invariant properties of anisotropic sheets. *Tappi* **70**, 117

Bradley S A, Carr S H J 1976 Mechanical loss processes in polysaccharides. *Polym. Sci.; Polym. Phys.* **14**, 111–24

Byrd V 1972a Effect of relative humidity changes during creep on handsheet paper properties. *Tappi* **55**, 247–52

Byrd V 1972b Effect of relative humidity changes on compressive creep response of paper. *Tappi* **55**, 1612–3

Caulfield D F 1985 A chemical kinetics approach to the duration-of-load problem in wood. *Wood Fiber Sci*. **17**, 504–21

Caulfield D 1980 Interactions at the cellulose-water interface. In: *Paper Science and Technology–The Cutting Edge*. Institute of Paper Chemistry, Appleton, WI

Caulfield D 1994 A thermodynamic basis for the 'ratchet' mechanism in cyclic humidity creep. In: Felless C, Langenberg T (eds.) *Proc. Moisture-induced Creep Behaviour of Paper and Board*. STFI, Stockholm, Sweden, pp. 89–103

Caulfield D, Weatherwax R C 1978 Tensile modulus of wet stiffened paper. In: Fundamental Research Committee (ed.) *Fiber Water Interactions in Papermaking*. Technical Division BPBIF, London, pp. 741–57

Caulfield D, Zauscher S 1997 Relaxation phenomena in Hydrogen-bond-dominated solids. In: Chalmers I R (ed.) *Proc. 3rd Int. Symp. Moisture and Creep Effects on Paper, Board and Containers*. Rotorua, New Zealand, pp. 17–37

Chatterjee S G, Ramarao B V 1997 Hysteresis in water vapor sorption equilibria of bleached kraft paperboard. In: Chalmers I R (ed.) *Proc. 3rd Int. Symp. Moisture and Creep Effects on Paper, Board and Containers*. Rotorua, New Zealand, pp. 121–34

Coleman B 1956 Application of the theory of absolute reaction rates to the creep failure of polymeric filaments. *J. Poly. Sci.* **20**, 447–55

Crook D M, Bennett W E 1962 Effect of humidity and temperature on the physical properties of paper. *British Paper and Board Industry Research Association*

de Ruvo A, Lundberg R 1972 *Lifetime of Paperboard at Constant Loading and Varying Climate* (in Swedish). STFI Rep. No. B131, Stockholm, Sweden

de Ruvo A, Lundberg, Martin-Löf R S, Söremark C 1976 Influence of temperature and humidity on the elastic and expansional properties of paper. In: Bolam F (ed.) *Fundamental Properties of Paper Related to Its Uses*. Technical Division BPBIE, London, Vol. 2, pp. 785–806

Dodson C T J, Herdman P T 1982 Mechanical properties of paper. In: Rance H F (ed.) *Handbook of Paper Science*. Elsevier, New York

Fellers C, Laufenberg T (eds.) 1994 *Proc. Moisture-induced Creep Behaviour of Paper and Board*. STFI, Stockholm, Sweden

Goring D A I 1966 Thermal softening, adhesive properties and glass transitions in lignin, hemicellulose and cellulose. In: Bolam F (ed.) *Consolidation of the Paper Web*. Clowes, London

Grossman P U A 1976 Requirements for a model that exhibits mechano-sorptive behavior. *Wood Sci. Technol*. **10**, 163–8

Gunderson D 1981 A method for compressive creep testing of paperboard. *Tappi* **64**, 67–71

Guthrie J B, Fulmer G E 1969 Characterization of saturated cellulosic webs by the creep failure method. *Tappi* **52**, 2181–290

Holliday L, White J W 1971 Stiffness of polymers in relation to their structure. *Pure Appl. Chem*. **26**, 545–82

Htun M, de Ruvo A, Fellers C J 1985 The dynamic mechanical behaviour of paper during drying. *Appl. Poly. Sci*. **30**, 597–604

Kestin J 1978 *A Course in Thermodynamics*. McGraw Hill, New York

Klason C, Kubàt J, Svensk 1976 Thermal transitions in cellulose. *Papperstidning* **15**, 494–500

Kolseth P, de Ruvo A 1983 The measurement of viscoelastic behavior for characterization of time–temperature–humidity dependent properties. In: Mark R E (ed.) *Handbook of Physical and Mechanical Testing of Paper and Paperboard*. Vol. 2, Marcel Dekker, New York

Kubàt J 1969 Rheology of paper. In: Eirich F R (ed.) *Rheology*. Academic Press, New York

Kubàt J, Lindbergson B J 1965 Damping transients in polymers during sorption and desorption. *Appl. Poly. Sci.* **9**, 2651–4

Kubàt J, Martin-Löf S, Söremark, Svensk C 1969 Thermal expansivity and elasticity of paper and board. *Papperstidning* **23**, 763–7

Meixner J 1960 Relaxationserscheinungen und ihre thermodynamische Behandlung. *Naturkunde* **26**, 259

Nielsen L E 1974 *Mechanical Properties of Composites and Polymers*. Marcel Dekker, New York

Nissan A H 1957 Rheological behavior of hydrogen bonded solids. I. Primary considerations. *Trans. Faraday Soc.* **53**, 700–9

Nissan A H 1967 The significance of hydrogen bonding at the surfaces of cellulose network structure. In: Marchessault R H, Skaar C (eds.) *Surfaces and Coatings Related to Paper and Wood*. Syracuse University Press, Syracuse, New York

Nissan A H 1977 *Lectures on Fiber Science in Paper, Pulp and Paper Technology*. Series No. 4, Joint Textbook Committee TAPPI CPPA

Nissan A H, Byrd V, Batten G Jr, Ogden R W 1985 Paper as an H-bond dominated solid in the elastic and plastic regimes. *Tappi* **68**, 118–24

Nissan A H, Sternstein S S 1962 Cellulose as a viscoelastic material. *Pure Appl. Chem.* **5**, 131–46

Nissan A H, Sternstein S S 1964 Cellulose–fiber bonding. *Tappi* **47**, 1–6

Nordvall E 1992 Ultrajudmätning av dragstyvhetsindex. Thesis, Institute of Paper Technology, KTH, Stockholm, Sweden

Onsager L 1931 *Phys. Rev*. Reciprocal relations in irreversible processes. **I. 37**, 405–26, Reciprocal relations in irreversible processes. **II 38**, 2265–79

Page D 1969 A theory for the tensile strength of paper. *Tappi* **52**, 671–81

Prigogine I, DeFay R 1954 *Chemical Thermodynamics*. Longmans Green, London

Rance H 1950 Fibre assemblages: the rheological behaviour of paper. In: Harrison V G W (ed.) *Some Recent Developments in Rheology*. United Trade Press, London, pp. 90–104

Rance H F 1953 The mechanical properties of paper. In: Meredith R (ed.) *Mechanical Properties of Wood and Paper*. North Holland, New York

Salmèn L, Back E L 1977 Simple stress–strain measurements on dy papers from 25 °C to 250 °C. *Svensk. Paperstidn.* **80**, 178–83

Salmèn L, Back E 1980 Moisture-dependent thermal softening of paper evaluated by its elastic modulus. *Tappi* **63**, 117–20

Steenberg B 1947 Paper as a visco-elastic body. *Svensk. Paperstidn.* **50**, 127–40

Sternstein S S 1961 The viscoelasticity of hydrogen-bonded materials. Thesis, Rensselear Polytechnic Institute, Troy, New York

Tobolsky A V 1960 *Properties and Structure of Polymers*. Wiley, New York

Wahba M 1965 Is the thermal transition of cellulose of the glass type?. *J. Textile Inst. Trans.* **56**, T218–21

Wink W A 1961 The effect of relative humidity and temperature on paper properties. *Tappi* **44**, 171A–8A

Zauscher S, Caulfield D, Nissan A H 1996 The influence of water on the elastic modulus of paper. Part 1. *Tappi* **79**, 178–82

Zauscher S, Caulfield D, Nissan A H 1997 Influence of water on the elastic modulus of paper. Part II. *Tappi* **80**, 214–3

Zhurkov S N 1965 Kinetic concept of the strength of solids. *Int. J. Fracture Mechan.* **1**, 311–23

D. F. Caulfield[a] and A. H. Nissan[b†]
[a] *USDA Forest Service, Madison, Wisconsin, USA*
[b] *Scarsdale, New York, USA*

Paper: Pulping and Bleaching

Wood pulping processes are broadly categorized as mechanical or chemical according to the relative importance of mechanical treatments compared to chemical modification, degradation, and/or solution of wood components.

The properties of wood pulps and papers made from such pulps are affected by the selection of wood species (see *Pulp and Paper: Wood Sources*#) and by the complex mechanical, chemical, and thermal processes occurring during pulping and bleaching. The physical characteristics of the fibers, their chemical composition, and the supramolecular properties of the constituent polymers can all change in these processes.

Reductions in air and water emissions from pulping and bleaching processes are being accomplished by process selection, increased process closure, effective recovery of chemicals, and/or end-of-pipe treatments.

1. Mechanical Pulping

Mechanical processes use attrition to separate wood into fibers and fiber fragments. The resulting pulp properties are highly dependent upon the applied energy and the nature of the physical interactions with the wood. Mechanical pulping processes use large quantities of electrical energy. Equipment design, chemical applications, and thermal energy are applied to reduce this energy requirement (Cannell 1999). Mechanical pulps are produced at high yields, often exceeding 90% of wood mass. The fibers tend to be stiff and produce bulky paper structures. Papers can also be made with good printing characteristics and opacity. Mechanical pulps are well suited for production of newsprint, publication papers, and some coated catalog and publication grades. Mechanical pulps are also used in tissue papers, cartons, and building boards.

1.1 Mechanical Pulping Processes

In groundwood (GW) mechanical pulping, wood logs or bolts are pressed against a revolving grinding stone, and fibers and fiber fragments are separated by abrasion. Pulp characteristics are influenced by the stone surface characteristics, stone surface speed, energy consumption, temperature, and pulp consistency in the grinder. Higher strength, longer-fibered pulps can be made by operating the grinding process at elevated temperatures (130 °C) and pressures up to 0.5 MPa in the pressurized groundwood (PGW) process.

Production of pulp by abrasion of wood chips or sawdust in a refiner between rotating steel disks allows greater flexibility in raw material selection and provides options for producing pulps with a wider range of properties. Refiner mechanical pulps (RMP) are produced at atmospheric pressures. In thermomechanical pulping (TMP), the refining process is steam pressurized to allow temperatures up to 170 °C. The wood lignin is softened and the wood fracture location is controlled, by temperature, to inside the fiber wall at 130 °C or in the lignin-rich middle lamella region between fibers at higher temperatures. Refiner mechanical pulping processes can utilize two stages in series to achieve the desired pulp qualities.

Small amounts of chemicals, such as sodium hydroxide, carbonate, sulfite, bisulfite, or combinations thereof, are sometimes added to produce chemimechanical (CMP) or chemithermomechanical (CTMP) pulps. Wood components are chemically modified by the additives to be more hydrophilic and, in some cases, are partly dissolved. Swelling of the wood is increased and the response to the mechanical pulping process is altered dramatically. Chemical additions affect the location of the fracture in the fiber structure as well as the nature of the fiber surfaces formed. Both of these factors influence the capacity of the fibers to develop strong bonds in the final paper product. Typical yields in chemimechanical pulping are in the 80–95% range.

Semichemical pulps are produced at yields between 65% and 80% by adding larger amounts of chemicals and/or increasing the pretreatment time or temperature. As larger amounts of chemicals are used, recovery of the spent chemicals becomes an economic and environmental consideration.

1.2 Bleaching Mechanical Pulps

Production of high brightness mechanical pulps requires consideration of the entire process including: selection of the wood source, in-process storage of the wood, the pulping process, and pulp bleaching, as well as the intended use of the product. Light-colored woods without fungal discoloration or dark-colored resins are preferred. Wood lignin and resins are modified and partially removed in the preparation of mechanical or semichemical pulps. Remaining lignin, resins, extractives, and other minor components of the wood contribute to the color of these pulps. Chemicals applied in pulping can reduce loss of brightness in the process and, in some cases, improve the color of the pulp. Mechanical pulp bleaching processes are designed to remove chromophoric structures and brighten the pulp with minimum losses of mass. Lignin-preserving bleaching is accomplished commercially primarily with hydrogen peroxide and/or sodium dithionate. Sequential bleaching stages with these chemicals are used to achieve high brightness. Chelating or complexing agents can be applied to control trace amounts of metal ions which can form colored complexes or adversely impact the bleaching process by decomposing hydrogen peroxide.

While reasonably high brightness can be developed with mechanical and chemimechanical pulps, it is difficult to prevent reversion or brightness loss over time. Exposure to heat and light accelerates brightness reversion. Brightness reversion often limits mechanical pulps to products with relatively short life cycles.

1.3 Environmental Emissions from Mechanical Pulping

Water-soluble wood components, pulping and bleaching chemicals, and wood substances dissolved in the mechanical pulping processes contribute biological oxygen demand (BOD), chemical oxygen demand (COD), and color to the aqueous effluent. Emissions of volatile organic compounds can occur in pulping and bleaching. Aqueous effluents are treated with primary and secondary treatment processes before discharge. “Zero effluent” production of bleached CTMP has been commercialized in Canada to avoid aqueous emissions (Meadows 1996).

2. Chemical Pulping Technology

Chemical pulping processes produce low brightness pulps in yields ranging from 40% to 65% by dissolving lignin, hemicellulose and cellulose carbohydrates, and extractives from the wood. At the lower yields, intact, individual fibers are readily separated from the wood structure. Chemical pulps made at the higher yield levels and containing larger amounts of the wood lignin require appreciable mechanical energy input from disk refining equipment to complement the chemical pulping process.

2.1 Alkaline Pulping

In the commercially dominant kraft (sulfate) process, sulfide and hydroxide ions are used to depolymerize and dissolve the wood extractives, lignin, and some hemicelluloses. The pulping reactions require one to three hours at temperatures up to 170 °C. Both batch and continuous pressurized digesters are used. Kraft pulping has been improved by the use of additives, selective control of chemical concentrations, and optimization of temperature profiles. Modified kraft pulping processes apply four principles to improve the selectivity of lignin removal: leveled-out hydroxide concentration, high initial hydrosulfide concentration, low sodium ion and dissolved lignin concentrations in the final part of the cook, and avoidance of high temperature. These principles have been employed in various ways to achieve improved pulping and the ability to pulp to lower lignin contents (Teder *et al.* 1997).

Anthraquinone (AQ) addition to kraft pulping increases the rate of the delignification reaction and the yield of carbohydrates. Although AQ is used at very low levels, less than 0.1% on wood, it is a relatively expensive chemical and is often used on an as-needed basis to provide incremental production capability.

Polysulfide (PS) ions in kraft pulping stabilize the carbohydrates against the alkaline degradation reaction known as "peeling." With PS addition, the yield of kraft pulp can be increased by 1–3%. To maintain the sulfur balance in the pulping-recovery cycle, it is necessary to regenerate PS from the recovered pulping chemicals. Polysulfide and anthraquinone can be applied together in the PS-AQ process to get synergistic pulp yield benefits.

Soda pulping using sodium hydroxide avoids the odorous sulfur-containing gases associated with kraft pulping. However, pulp strength properties are adversely impacted by the high concentrations of hydroxide ion needed to achieve practical pulping rates. AQ can be added for improved pulping rates.

Alkaline sulfite pulping processes have shown promise for increased pulp yield and good pulp quality. AQ and/or methanol are additives in some alternatives. Commercial application has been limited by the lack of a cost-effective recovery option.

2.2 *Sulfite and Solvent Pulping Processes*

In sulfite pulping, usually conducted under acid or neutral conditions, sulfur dioxide and/or bisulfite ions react with lignin to produce water-soluble sulfonated lignins that are degraded by acid hydrolysis reactions. Sulfite pulping can be controlled to produce pulps over a wide range of lignin contents and pulp yields. The residual lignin in sulfite pulps is sulfonated and is relatively easy to remove in pulp bleaching processes. The conditions of sulfite pulping lead to extensive hydrolysis of the carbohydrates, and sulfite pulps typically have less strength potential than kraft pulps in papermaking applications. The recovery and regeneration of chemicals in the waste liquors from sulfite pulping are more difficult than for kraft liquors.

Solvent pulping processes have been extensively evaluated in laboratory and pilot tests, but have found only limited commercial application. Many organic solvents have been used including alcohols, acids, ketones, esters, and others. Often an inorganic acid is added to promote hydrolytic cleavage of lignin bonds. Solvent pulping offers possibilities for alternative chemical recovery processes with lower capital costs and by-product recovery. Very high solvent recovery must be achieved for economic and environmental reasons.

2.3 *Recovery of Chemicals for Pulping*

Efficient recovery and regeneration of pulping chemicals are essential for chemical pulping processes. Wood components dissolved in the pulping process are converted to useful by-products, such as tall oil soap and lignin, or to energy by combustion. The availability of a cost-effective chemical and energy recovery cycle for kraft pulping contributes to the dominance of this process. Kraft pulping operations are often energy self-sufficient.

In the kraft pulping/recovery cycle, pulp from the digester system is washed with water to separate a solution of spent pulping chemicals and dissolved wood components known as black liquor. This black liquor is concentrated by multiple effect evaporation to a high solids concentration (>50%) and then burned in a heat recovery boiler. The inorganic compounds are removed from the furnace as a molten smelt and are dissolved in water to form green liquor. The active pulping liquor, a solution of sodium hydroxide and sodium sulfide called white liquor, is regenerated by causticizing the green liquor with lime. The lime is also regenerated in a separate cycle using a kiln to convert calcium carbonate to calcium oxide.

The recovery cycle operations for kraft pulping are capital intensive, and many alternatives have been evaluated. Gasification processes under development for burning the concentrated black liquor offer promise for improved energy recovery by application of combined cycle power generation.

2.4 *Bleaching Chemical Pulps*

Chemical pulps contain 2–5% lignin after the pulping process and are dark brown in color. For white products, these chromophoric residues are removed in multistage processes with oxidative chemicals such as oxygen, chlorine, hypochlorites, chlorine dioxide, and hydrogen peroxide. Alkaline extraction facilitates removal by solubilization of the oxidatively degraded materials.

An oxygen stage is often used as a transition between pulping and bleaching. Oxygen delignification removes 30–60% of the residual lignin. Oxygen stages operate at either medium (~13%) or high (~30%) consistency with 1–4% sodium hydroxide addition. Temperatures range from 80 to 120 °C with oxygen pressures up to 10 atmospheres. Two oxygen stages in series are sometimes used. Oxygen stages are integrated with the washing of the unbleached pulp so that wood components dissolved in this stage can be processed in the chemical recovery cycle. This closure of the process is a major advantage of oxygen compared to many other bleaching chemicals where the effluent is not recovered and must be treated before discharge to the environment.

The chemistry of pulp bleaching has changed dramatically since the mid-1980s, when trace amounts of dioxin were found in bleached chemical pulps and effluents from bleaching processes using chlorine in the first oxidative bleaching stage at low pH. Dioxin formation was avoided by changing chlorination stage conditions or by bleaching without chlorine (elemental chlorine free, ECF). Bleaching technologies also were developed that avoid all chlorine bleaching chemicals (totally chlorine free, TCF) (Nelson 1998). ECF bleaching processes provide several pulp quality advantages relative to TCF. Comparative evaluations demonstrated that treated effluents from the two approaches are comparable. ECF bleaching sequences have been preferred by the industry for kraft pulp bleaching.

Hypochlorite bleaching uses chlorine at alkaline pH conditions with either sodium or calcium as the base. Hypochlorite reactions with lignin produce chloroform as a by-product. As a result, hypochlorite bleaching has been largely replaced by hydrogen peroxide or by increased use of chlorine dioxide.

Chlorine dioxide is a very effective pulp bleaching chemical. Initially, because of its high cost, chlorine dioxide was used mostly in the later stages of bleaching to remove the last traces of residual color in the pulp. In ECF bleaching, chlorine dioxide is used in the first stage as a delignifying chemical to replace chlorine. Chlorine dioxide is a very selective chemical that removes lignin residues with very little chemical attack on the carbohydrates in the fiber. The first dioxide stage of a bleaching sequence is usually followed by an alkaline extraction stage where oxidized fragments of lignin are solubilized and removed from the pulp. Alkaline extraction stages are sometimes used between two dioxide stages in the last stages of bleaching. Recycling of ECF bleaching effluents to recovery requires additional processing to remove sodium chloride that can accumulate in the pulping–recovery cycle liquors and cause corrosion problems.

Hydrogen peroxide has become more important in both ECF and TCF pulp bleaching processes. Hydrogen peroxide is most effectively used as a pulp brightening chemical. It is often used as an additive to alkaline extraction stages.

TCF bleaching applies combinations of oxygen, ozone, and peroxide bleaching stages. Trace metal ions present in the unbleached pulp can adversely affect the stability of peroxide resulting in wasteful loss of the chemical as well as radical species that can attack the pulp carbohydrates. The effects of trace metals can be reduced by removal of the metals with an acid or chelant treatment.

Ozone was developed as a commercial bleaching chemical in the late 1980s in response to the need for more effective chlorine-free bleaching chemicals. Ozone stages can be operated at medium or high consistency. Ozone is a very reactive chemical, and undesired reactions with the pulp carbohydrates must be controlled. Efficient mixing of the ozone gas with the pulp is a major challenge. Developments in the technology for generation of ozone have resulted in ozone becoming a cost-competitive option for the industry. Effluents from ozone stages early in the process can be recycled back to the chemical recovery cycle.

Achieving very high brightness levels has been a challenge with TCF bleaching processes. To facilitate these processes, it is common to extend delignification of unbleached pulp to lower lignin contents and to operate the oxygen stage at higher delignification levels. These steps, as well as the alkaline conditions and high temperatures used in TCF peroxide stages, raise concerns about the loss of additional carbohydrates and thus pulp yield.

The most significant effects of bleaching on the properties of chemical pulps are related to the extent of oxidative chemical attack by the bleaching chemicals on the cellulose and hemicellulose polymers. Some hemicelluloses can be dissolved during bleaching; if significant amounts are intentionally removed, as in the production of high-purity α-cellulose pulps, the properties of the pulp can be strongly affected.

The oxidative chemicals used in pulp bleaching have different potentials for affecting pulp properties. These chemicals can both depolymerize the hemicellulose and cellulose polymers and introduce new functional groups into these polymers. Some depolymerization can occur without strength loss but, beyond a certain point, strength properties decrease sharply as the polymer molecular weight decreases. Depolymerization is often monitored by empirical measurements of the solution viscosity of the pulp in selected solvents, but changes in the chemical composition of the pulp also affect the results of this test.

By proper selection of bleaching chemicals and operating conditions, it is possible to minimize changes in pulp strength properties during bleaching. Usually, the most active chemicals, such as oxygen and ozone, are applied in the early stages of bleaching, where the larger amounts of residual lignin present in the pulp rapidly consume the chemical and protect the carbohydrates. Hypochlorites, when used, are applied in intermediate stages of bleaching. The amount of undesired chemical attack by hypochlorites is reduced by controlling the reaction pH and by limiting the amount of the chemical used. Hydrogen peroxide bleaching can be a relatively mild process and is often applied in the final bleaching stages.

Chlorine dioxide has the least effect on pulp strength properties and is favored for the final stage of bleaching when pulps having high brightness and high strength are desired.

3. Biotechnology in Pulping and Bleaching

Application of biotechnology in pulping and bleaching has been evaluated extensively for both chemical and mechanical pulp processes. A commercial product uses a fungus to degrade extractives in wood chips before pulping to prevent later deposition of pitch on process equipment. Biomechanical pulping uses lignin-degrading microorganisms, primarily white-rot fungi, to pretreat wood chips prior to refining. Significant reductions in energy requirements and increases in strength properties have been reported, with biomass losses up to 25%. Many fungi strains have been screened to take advantage of the natural differences in the lignin- and carbohydrate-degrading ability of these organisms. Fungal pretreatments have also been evaluated for improved delignification and bleachability in sulfite and kraft pulping (Messner 1998).

Isolated xylanase enzymes are used to enhance chemical pulp bleaching. Xylanase treatments before bleaching result in higher brightness after bleaching or lower bleaching chemical costs.

4. Process Effects on Pulp Properties

The physical properties of chemical pulps depend markedly on the pulping yield. The strength of wood pulp fibers is primarily caused by the high-molecular weight, highly oriented cellulose polymer in the fiber. Since the cellulose component is less affected by the pulping processes, the fraction of cellulose in the pulp increases as the yield decreases. Also, the number of fibers per unit weight of pulp increases as more material is removed by pulping. Pulp strength properties thus improve as yield decreases, up to a point where chemical attack on the cellulose becomes significant.

In addition to the basic fiber strength, hydrogen bonds formed between the fiber elements in the paper affect the strength developed from a pulp. The development of hydrogen bonds in paper is determined to a great extent by how the fiber swells in water during the preparation of the fibers for papermaking. Many of the effects of pulping and bleaching on pulp properties can be understood in terms of the relative amounts of the hydrophobic lignin polymer removed and the hydrophilic hemicellulose retained in the pulp fiber. The location in the fiber structure and degree of crystallinity of hemicellulose can also be affected by the pulping process.

Mechanical action during processing of pulps, such as occurs when pulps are discharged by hot blowing from high-pressure pulping vessels, pumped, mixed, or conveyed, can decrease fiber strength. Such mechanical action may cause undesired ‘‘dislocations’’ or weak spots in the fiber wall or formation of fiber debris by mechanically disrupting the basic fiber structure.

5. Conclusions

Many alternatives for production of pulps with wide ranges of end-use specific properties have been developed and are applied commercially. Environmental issues are being addressed with new technologies, alternative chemicals, and progressive reduction of process emissions.

See also: Cellulose: Chemistry and Technology#; Wood, Constituents of*; Pulp and Paper: Wood Sources#

Bibliography

Akhtar M, Blanchette R A, Myers G, Kirk T K 1998 An overview of biomechanical pulping research. In: Young R A, Akhtar M (eds.) *Environmentally Friendly Technologies for the Pulp and Paper Industry*. Wiley, New York, Chap. 10, p. 309

Axegard P 1998 Minimal impact bleached kraft pulp mill. *Pulp Pap. Can.* **99** (4), 81–6

Axegard P, Carey J, Folke J, Gleadow P, Gullichsen J, Pryke D C, Reeve D W, Swan B, Uloth V 1997 Minimum-impact mills: issues and challenges. In: *TAPPI Proceedings Environmental Conference & Exhibition*. TAPPI Press, pp. 529–41

Cannell E 1999 Mechanical pulping technologies focus on reducing refining energy. *Pulp Pap.* **73** (5), 63–70

Dence C W, Reeve D W 1996 *Pulp Bleaching—Principles and Practice*. TAPPI Press, Atlanta, GA

Green R P, Hough G 1992 *Chemical Recovery in the Alkaline Pulping Processes,* rev. edn. TAPPI Press, Atlanta, GA

Hergert H L 1998 Developments in organosolv pulping—an overview. In: Young R A, Akhtar M (eds.) *Environmentally Friendly Technologies for the Pulp and Paper Industry*. Wiley, New York, Chap. 1, p. 5

Hynninen P 1998 Environmental control. In: Gullichsen J, Paulapuro H (eds.) *Papermaking Science and Technology*, Book 19 Fapet Oy, Helsinki, Finland, Chap. 4, p. 31

Kocurek M J, Grace T M, Malcolm E 1989 Alkaline pulping. In: Joint Textbook Committee of the Paper Industry (eds.) *Pulp and Paper Manufacture*. TAPPI Press, Atlanta, GA, Vol. 4

Kocurek M J, Ingruber O V, Wong A 1985 Sulfite science and technology. In: Joint Textbook Committee of the Paper Industry (eds.) *Pulp and Paper Manufacture*. TAPPI Press, Atlanta, GA, VOl. 4

Kocurek M J, Leask R A 1987 Mechanical pulping. In: Joint Textbook Committee of the Paper Industry (eds.) *Pulp and Paper Manufacture*. Vol. 2. TAPPI Press, Atlanta, GA

McDonough T J 1995 Recent advances in bleached chemical pulp manufacturing technology: extended delignification, oxygen delignification, enzyme applications, and ECF and TCF bleaching. *TAPPI J.* **78** (3), 55–62

Meadows D G 1996 Meadow Lake (Saskatchewan, Millar Western's pulp mill) makes fourth year of zero effluent pulping. *TAPPI J.* **79** (1), 63–8

Messner K, Koller K, Wall M B, Akhtar M, Scott G M 1998 Fungal treatment of wood chips for chemical pulping. In: Young R A, Akhtar M (eds.) *Environmentally Friendly Technologies for the Pulp and Paper Industry*. Wiley, New York, Chap. 12, p. 385

Nelson P J 1998 Elemental chlorine free (ECF) and totally chlorine free (TCF) bleaching of pulps. In: Young R A, Akhtar M (eds.) *Environmentally Friendly Technologies for the Pulp and Paper Industry*. Wiley, New York, Chap. 7, p. 215

Sundholm J 1998 Mechanical pulping. In: Gullichsen J, Paulapuro H (eds.) *Papermaking Science and Technology*, Book 5. Fapet Oy, Helsinki, Finland

Teder A, Lindgren C, Lindstrom M 1997 Basic principles for the next generation of selective modified kraft pulping. *Papier*. **51**(6A), 42–46

H. L. Hintz
Westvaco Corporation, New York, USA

Paper: Strength and Stiffness

Paper is the material of choice for a host of products such as books, graphic products, packaging materials, shipping containers, and personal care products. In part, paper is used because the physical properties can be tailored to fit the end-use in the most economical way. For example, paper can be produced with high stiffness and strength for use in a box or high compliance for tissue paper. Examples of standards used in the paper industry are ISO and national standards such as TAPPI and SCAN.

1. Definitions

Consider a uniaxial tensile test of a paper test specimen that yields as shown in a stress-strain diagram (Fig. 1). For tensile testing of paper, emphasis must be placed on the grips to minimize slippage. In standard practice, rectangular test pieces are used, often with special line-load clamps.

The compressive strength of paper is a more ambiguous measurement, with the results dependent on the test method. The obstacle in compression is the low buckling resistance of a sheet of paper. Several compressive strength tests are commonly used including ring crush and short-span compressive strength. To evaluate the entire stress–strain curve in compression, special equipment is used (see *Paper: Creep**).

For tensile properties, the in-plane stiffness is evaluated from the initial slope of the curve. The paper industry often utilizes the terms tensile stiffness index E^w to characterize in-plane stiffness and tensile index σ_T^w to characterize in-plane strength. These properties equal force divided by test piece width b and grammage w (mass per unit area). Alternatively, tensile stiffness E^b and tensile strength σ_T^b are used, with dimensions in force per unit width. The practical reason for using these properties is that paper is judged by its properties per unit mass rather per unit width and thickness (cross-sectional area). Furthermore, the rough surface and low stiffness in the thickness direction produces thicknesses that are highly dependent on the techniques employed. By this approach, thickness measurements are avoided.

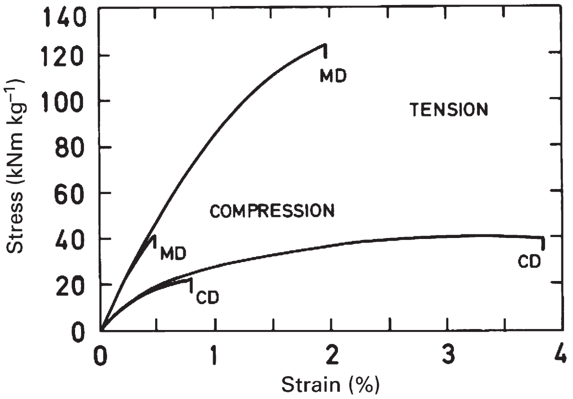


Figure 1
Stress-strain curves for paper in tension and compression. MD is the machine direction. CD is the cross machine direction. (Based on results by Fellers and Norman 1996.)

Tensile stiffness index E^w and tensile index σ_T^w are related to elastic modulus E and tensile failure stress σ_T, each divided by density ρ:

$$E^w = \frac{E}{\rho}\ (\mathrm{N\,m\,kg^{-1}}) \tag{1}$$

$$\sigma_\mathrm{T}^w = \frac{\sigma_T}{\rho}\ (\mathrm{N\,m\,kg^{-1}}) \tag{2}$$

These terms are useful when comparing properties of different materials. In many publications, tensile strength is expressed as breaking length σ_T^l, which is equal to the length of paper that would break under its own mass. The relationship between breaking length and tensile index is given in Eqn. (3), where g is the acceleration due to gravity, $9.81\,\mathrm{m\,s^{-2}}$:

$$\sigma_\mathrm{T}^l = \frac{\sigma_\mathrm{T}^w}{g}\ [\mathrm{m}] \tag{3}$$

For compressive properties, the in-plane stiffness is the same as in tension. In order to avoid the use of two terms for the same property, the term tensile stiffness is often used instead of compressive stiffness. The compression index σ_C^w is on the order of 30% of the tensile index.

The bending stiffness per unit width S^b is related to elastic modulus E and the thickness t by:

$$S^b = \frac{Et^3}{12} = \frac{E^w w^3}{12\rho^2} \,(\mathrm{N\,m}) \tag{4}$$

Bending stiffness index S^w, stiffness related to the grammage of the material, consequently may be defined as:

$$S^w = \frac{S^b}{w^3} \,[(\mathrm{N\,m})^7 \mathrm{kg}^{-3}] \tag{5}$$

Table 1 gives examples of how different materials may be ranked based on force per cross sectional area (stress) or based on properties per width and mass (index). If materials are ranked based on stress, steel and glass are much stiffer and stronger in the plane of the sheet material than are paper and wood. If properties are based on index, the in-plane properties are in the same order for all four materials. The great difference appears when comparing the bending stiffness at equal mass of the materials, where much higher values are achieved for paper and wood.

There are many other different mechanical properties that are routinely measured for paper and paper products including other in-plane and out-of-plane properties; compressive properties, burst strength, and tearing strength. The amount of information that has been recorded on this subject is extensive and the interested reader is referred to the bibliography for in-depth discussions of the mechanical properties of paper. The discussion below will be restricted to the relationships among stiffness, tensile, and compression properties; and effects of the paper structure; and the papermaking process.

2. *Orthotropic Nature of Paper*

The mechanical properties of machine-made papers exhibit anisotropy, and the assumption of orthotropy appears to be reasonable. Figure 1 shows uniaxial stress strain curves for paper tested in the machine direction (MD) and the cross machine direction (CD) in tension and compression. Typically, the MD stiffness and strength are 1.5 to 3 times greater than the CD values, but much higher stiffness ratios can be achieved. Note that Fig. 1 also shows that paper is weaker in compression than tension. The directional dependence of properties is due to a combination of fiber orientation produced when the fibers are laid on the forming wire and MD restraint imposed during drying. Stiffness and strength in the thickness direction, ZD, is typically 50 to 100 times smaller than the in-plane stiffnesses. The Poisson ratios for the coupling between the in-plane and ZD directions are very close to zero. Typically, the two Poisson ratios for in-plane deformation have a geometric mean in the range of 0.1 to 0.4. Handsheets are typically isotropic with a 0.3 Poisson ratio.

Mann *et al.* (1980) have used ultrasonic techniques to measure the nine elastic constants that fully characterize the 3-D elastic behavior of paper. For in-plane properties, Baum *et al.* (1981) found that the in-plane shear modulus is 0.387 of the geometric mean of the two longitudinal stiffnesses. The implication of this relation is that paper is considered to have only three independent in-plane stiffness properties. For a given furnish, Htun and Fellers (1982) determined that the geometric mean of the two longitudinal moduli is approximately invariant with fiber orientation and drying restraint.

3. *Stiffness Measurements*

Standard mechanical techniques are considered the reference methods for determining stiffness properties of paper. Alternatively, nondestructive measurement

Table 1
Comparison of stiffness and strength indexes of different structural materials. (Gordon 1979, completed with data for paper by the authors.)

Material	Density ($\mathrm{kg\,m^{-3}}$)	Elastic modulus (GPa)	Tensile stiffness index ($\mathrm{MN\,m\,kg^{-1}}$)	Tensile failure stress (MPa)	Tensile index ($\mathrm{kN\,m\,kg^{-1}}$)	Bending stiffness index ($(\mathrm{N\,m})^7\,\mathrm{kg}^{-3}$))
Structural steel	7800	210	25	1500	192	0.04
Glass	2400	73	24	150	63	0.44
Wood (spruce) along the grain	500	14	25	103	206	9.33
Paper, linerboard, MD	700	9	13	75	107	2.19

of in-plane stiffnesses can be obtained by ultrasonic measurements. Since paper is a viscoelastic material, the ultrasonic modulus is typically greater than the stiffness measured from a tensile test. The much slower tensile test permits relaxation and creep to occur, therefore, the modulus measured in tension is a function of the strain rate, which is always specified.

The bending stiffness of paper is usually obtained using a two-point bending method or by strip resonance. In these tests, the level of loading must be kept small so that the material does not yield.

As seen from Eqn. (5), high bending stiffness of paper can be achieved by selecting fibers and papermaking processes to produce a low density structure as well as a high tensile stiffness index. For heavier grades, the paper is produced in several layers to optimize performance. High tensile stiffness is readily obtained by preventing shrinkage during drying.

4. Development of Strength in Paper

The strength of paper is affected by the fiber source, papermaking conditions, and environmental factors. The following discussion of these parameters will focus on tensile strength, however, the same general trends apply to other strength properties. Page (1969) developed a now classical equation that expresses the effect of the fiber and network structure on tensile index:

$$\frac{1}{\sigma_{\mathrm{T}}^{w}} = \frac{9}{8\sigma_{\mathrm{ZS}}^{w}} + \frac{12}{\tau_{\mathrm{S}} l \alpha} \tag{6}$$

where σ_{ZS}^{w} = zero span strength (a measurement of the fiber strength), τ_{S} = shear bond strength, l = fiber length, and α = relative bonded area per gram of fiber. The strength of a sheet will always be less than zero-span strength. The key components are the strength of the individual fibers, the shear strength of the bonds between fibers, and the degree of bonded area. If either bond strength or bonded area is low the sheet is bond-strength limited. If bond strength is high and bonded area adequate, the sheet is fiber-strength limited. Figure 2 illustrates the fracture zone for these two extreme cases.

The Page equation was developed for an isotropic sheet produced from straight fibers. If the fibers are very curly or have many defects or microcompressions that lowers the strength, the equation must be used with caution. Process variables can have a dramatic effect on sheet stiffness and strength. Figure 3 shows tensile and tensile stiffness index vs. density for a selected number of handsheets made from different pulps.

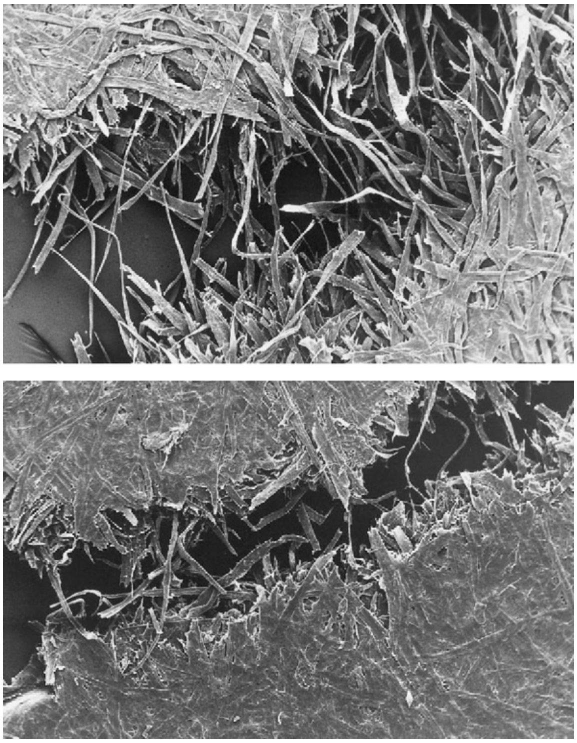


Figure 2
The fracture zones in papers subjected to tensile loading up to break. The upper figure shows a bond-strength-limited situation. The fibers are pulled out of the network with few fiber breaks. The lower figure shows a fiber-strength-limited situation. Few fibers are pulled out. Instead fiber fractures occur (after Fellers and Norman 1996).

Beating fibers increases fiber flexibility, which in turn increases the bonded area, producing a greater density and a stronger sheet. Pressing also increases the bonded area and produces similar strength improvement. Much of the strength of a machine-made sheet is developed in the drying process. MD tension in the web acts to restrain the sheet, keeping fibers straight and greatly increasing the strength. CD shrinkage produces a highly undesirable shrinkage gradient, causing gradients in strength and stiffness. It is possible to improve the smoothness of paper without significantly changing its mechanical properties by reducing the thickness and increasing the density by mechanical actions, such as calendering. Although calendering increases the density, it does not increase the bonded area, however, it has a tendency to reduce ZD strength, but with little effect on the in-plane properties.

Figure 5 illustrates the mechanism of failure in compression for paper, where strength is limited by the compressive properties of the individual fibers.

Compressive strength follows the same trends as tensile strength with papermaking factors, Fig. 4, but

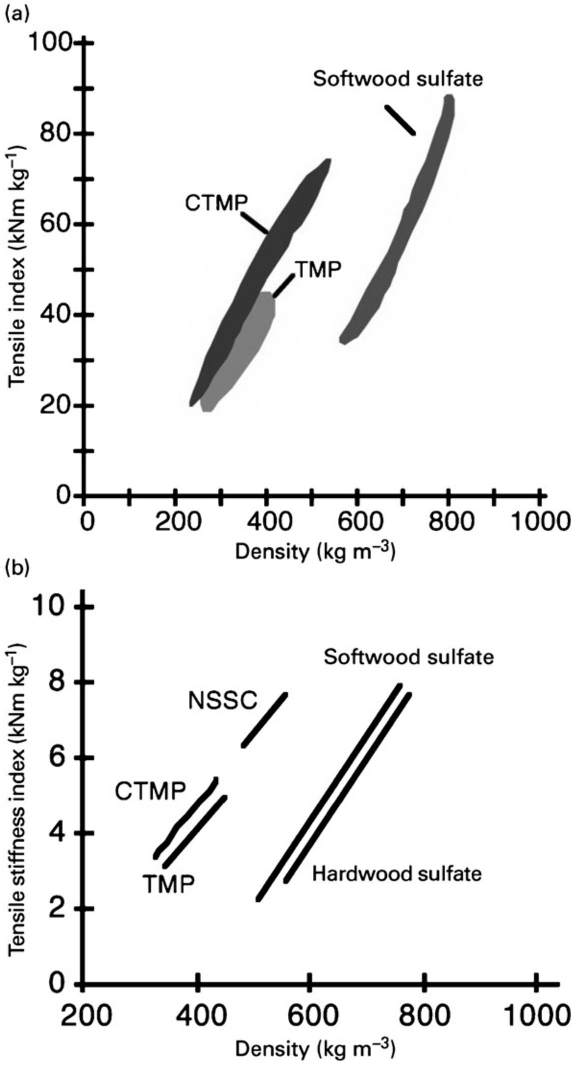


Figure 3
Tensile index and tensile stiffness index versus density for a selected number of handsheets made from different pulps (after Fellers and Norman 1996).

is limited by the compressive strength of the fibers, as shown in Fig. 5. Compressive strength is a fundamental property for the corrugated board components, linerboard, and corrugated medium.

5. *Environmental Effects*

All the mechanical properties of paper are highly dependent on the moisture content and temperature of the specimen, but moisture is by far the more important. As either the moisture content or temperature increases, the mechanical properties of the sheet decrease (see *Paper: Effects of Moisture and Temperature**). Therefore, it is very important that paper properties be reported at standard conditions of 50% relative humidity and 23 °C.

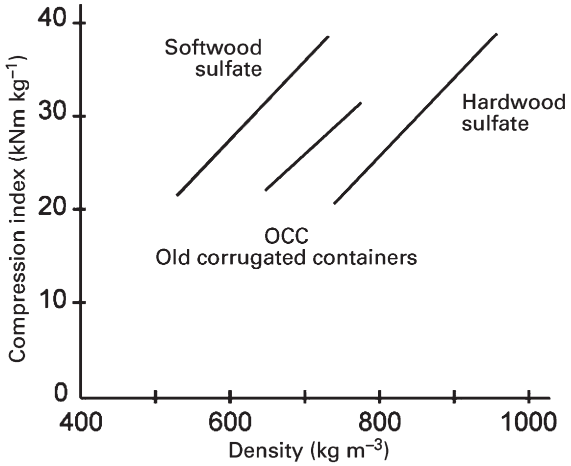


Figure 4
Compressive strength is a fundamental property for corrugated board components, linerboard and corrugated medium. Increased density by beating or pressing greatly improves the property (after Fellers and Norman 1996).

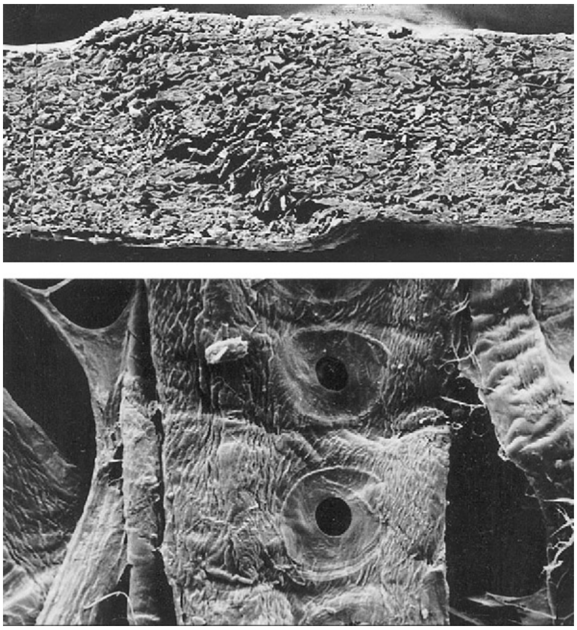


Figure 5
Compressive failure of paper. The strength is limited by the compressive strength of the individual fibers (after Fellers and Norman 1996).

6. Summary

There is a substantial amount of literature on strength and stiffness properties of paper, therefore, this article is an introduction to the subject. Stiffness and strength of paper are highly dependent on the source of fibers and the papermaking process, and a vast range of properties can be developed for paper, which is truly an engineered structure.

See also: Paper: Effects of Moisture and Temperature*

Bibliography

Baum G A 1993 Subfracture mechanical properties. In: Baker C F (ed.) *Trans. 10th Fundamental Research Symp. Products of Papermaking*. Oxford, UK. Pira International, Surrey, UK, pp. 1–126

Baum G A, Brennan D C, Habeger C C 1981 Orthotropic elastic constants of paper. *Tappi* **64**, 97–101

Bristow J A, Kolseth P 1986 *Paper Structure and Properties*. Dekker, New York

Fellers C, de Ruvo A, Htun M, Carlsson L, Engman C, Lundberg R 1983 *Carton Board—Profitable Use of Pulps and Processes*. Swedish Forest Products Research Laboratory

Fellers C, Norman B 1996 In: *Paper Technology*. KTH, Stockholm, pp. 283–325

Fellers C, Wellmar P, Kolseth P 1995 *Unified Symbols for Expressing Properties of Paper. Part 1. Mechanical Properties*, Report P 002. STFI, Stockholm, Sweden

Gordon J E 1979 *Structures or Why Things Don't Fall Down*. Penguin, New York

Haraldsson T, Fellers C, Kolseth P 1994 A method for measuring the creep and stress-strain properties of paperboard in compression. *J. Pulp Paper Sci.* **20**, J14–20

Htun M, Fellers C 1982 The invariant mechanical properties of oriented handsheets. *Tappi* **65**, 113–7

Kortshot M T 1993 The role of the fiber in the structural hierarchy of paper. In: Baker C F (eds.) *Trans. 10th Fundamental Research Symp. Fundamentals of Papermaking Materials*. Oxford, UK. Pira International, Surrey, UK, pp. 351–99

Mann R W, Baum G A, Habeger C C 1980 Determination of all nine orthotropic elastic constants for machine-made paper. *Tappi* **63**, 163–6

Mark R E 1983 *Handbook of Physical and Mechanical Testing of Paper and Paperboard, Vol. 1*. Dekker, New York

Niskanen K 1993 Strength and fracture of paper. In: Baker C F (eds.) *Trans. 10th Fundamental Research Symp. Products of Papermaking*. Oxford, UK. Pira International, Surrey, UK, pp. 641–726

Niskanen K 1998 *Paper Physics*. Fapet Oy, Helsinki, Finland

Page D 1969 A theory for the tensile strength of paper. *Tappi* **52**, 674–81

Wahlström T, Fellers C 1998 Biaxial straining of handsheets during drying— effect on in-plane mechanical properties. In: *International Engineering Conference, Book 3*. Tappi, Miami, FL,, pp. 1055–67

C. Fellers[a] and D. W. Coffin[b]

[a] *Skogsindustrins Tekniska Forskningsinstitut, Stockholm, Sweden*

[b] *Institute of Paper Science and Technology, Atlanta, Georgia, USA*

Paper: Structure

The mechanical behavior of a fiber network, such as a paper sheet composed of pulped wood fibers, depends on the fiber and fiber-to-fiber bond properties and the geometrical structure of the bonded fibrous network. The principal objective of a mathematical model is to predict the mechanical properties of the fiber network in terms of the properties of the fibers, bonds, and structure of the fiber network. In paper, the fibers are often collapsed, ribbon-like structures, which are bonded together primarily by hydrogen bonds formed when the sheet is pressed and subsequently dried.

The first attempt to develop a mathematical model for predicting the mechanical behavior of a fiber network such as paper was made by Cox (1952), who demonstrated how fiber orientation distribution influences the elastic behavior of the system. The analysis predicted all the elastic constants of a fiber mat in terms of the fiber orientation distribution expressed as a Fourier series.

Cox assumed that the fibers experience an axial strain corresponding to the sheet strain acting in the direction of the axis of the fiber. Cox's original theory has been subjected to numerous refinements and extensions which take into account the structure of the fiber network, the modeling of the fibers, the effects of drying-induced strains, the modeling of bond deformation, the nonlinear constitutive effects of the deformation response of the fibers and bonds, and the effect of fiber buckling for fibers that experience compressive strains.

The Cox model was modified by Qi (1997) to account for transverse normal and shearing components of strain in the fiber by assuming that the fiber takes on the same strains as those of the sheet. This assumption is equivalent to that used in laminate theory; therefore, if the fibers are assumed to lie in the plane of the sheet, the Qi model gives the same result as that predicted by laminate theory.

Early work regarding the development of mathematical models can be found in Algar (1966), Dodson (1973), Kallmes (1972a, b), and Van den Akker (1970, 1972). General aspects of paper physics can be found in books by Niskanen (1998) and by Deng and Dodson (1994). A review of research involving the modeling of the mechanical behavior of paper was given by Ramasubramanian and Wang (1999).

In this article, the basic assumptions and results of current mathematical models for predicting the in-plane mechanical behavior of paper are presented. The approach that is presented is based on the mesomechanics of the fibers that comprise a fiber network following the work of Perkins (1983, 1990).

1. Basic Elements of the Mathematical Model

The choice of the mesoelement that is used in the construction of a mathematical model depends on the

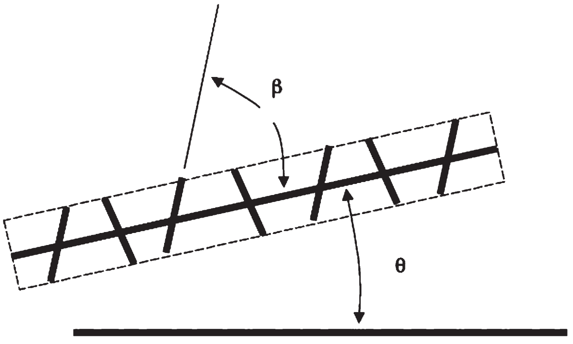


Figure 1
Typical fiber of length, λ_{f}, and orientation, θ, showing the angle of crossing fibers, β.

density of fiber sheet. In the case of low to medium density paper (to be defined subsequently), a typical mesoelement can be taken as an individual fiber along with the portions of crossing fibers that cross it and bond it to the fiber network. In order to develop a set of simple, closed-form expressions for the elastic moduli of the paper sheet, it is necessary to assume that the fibers are straight and that they lie in the plane of the paper. As discussed in the sequel, these assumptions can be relaxed, however, it is then necessary to perform the prediction of sheet properties numerically. Restricting attention to the low-density system comprised of straight fibers, the typical mesoelement of length, λ, and orientation, θ, is shown in Fig. 1. The mesoelements are coupled to the network by means of the crossing fibers. The strains in the sheet are presumed to be transmitted to the mesoelements by bending and shearing deformation of the crossing fibers and by shearing deformation of the fiber–fiber bonds. Thus the axial strain in the mesoelement fiber is not uniform but varies from the ends, where it is zero, to the middle, where its value is at its maximum. If the mesoelement is long enough and if the coupling is weak, the mesoelement strain is less than that associated with the sheet.

The model is further illustrated by Fig. 2, which shows a portion of a fiber that is coupled by two crossing fibers to the remainder of the network. The boundary between the element and the network is depicted by a dashed line. This boundary is assumed to be located a distance, $l/2$, from the centerline of the primary fiber, where l represents the center–center distance between bonds, l_{b} is the bond length along the fiber, and w_{f} and t_{f} are the width and thickness of the fiber, respectively. If l is small in comparison with w_{f}, as would be expected in moderately dense paper, the coupling is primarily attributable to the shearing deformations of the fiber–fiber bonds. In a very low-density system such as tissue paper, the bending and shearing deformation of the crossing fibers may be substantial.

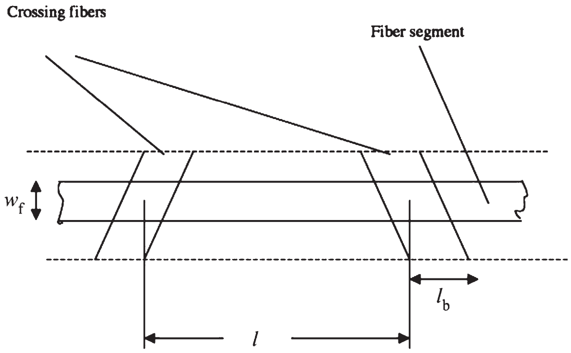


Figure 2
Portion of a fiber segment along a mesoelement showing the bond centroid distance, l, the bond length, l_{b}, and the fiber width, w_{f}. The mesoelement boundary applicable to low-density networks is shown as the dashed line.

On the assumption that the fiber is linearly elastic, the stress, σ_{f}, at any point along the fiber is related to the fiber strain, ε_{f}, at that point by

$$\sigma_{\mathrm{f}} = E_{\mathrm{af}}\varepsilon_{\mathrm{f}} \tag{1}$$

where E_{af} is the effective axial modulus of the fiber. The effective axial modulus depends not only on the inherent properties of the fiber but also on how the fiber is bonded to other fibers in the network. Thus, there is a stiffening effect due to the reinforcement of other bonded fibers. E_{af} can be considered as an inherent property of the paper system, or its value can be estimated in terms of an idealized paper network. For example, if the fibers are idealized (Figs. 1 and 2) as thin rectangular strips (e.g., perfectly collapsed thin-walled springwood fibers), and if the density of the network is low enough so that on average the fiber surface is only half (or less) in contact with other fibers in the network, then it can be shown that

$$E_{\mathrm{af}} = \frac{(l/l_{\mathrm{b}})(E_{\mathrm{fL}} + E_{\mathrm{fT}})}{C_{\mathrm{B}}\left(\dfrac{E_{\mathrm{fL}} + E_{\mathrm{fT}}}{E_{\mathrm{fL}}}\right)\left(\dfrac{l - l_{\mathrm{b}}}{l_{\mathrm{b}}}\right)} \tag{2}$$

where

$$C_{\mathrm{B}} = 1 + \frac{E_{\mathrm{fT}}}{E_{\mathrm{fL}}}\frac{\tanh B}{B} \tag{3}$$

and

$$B = \frac{l_{\mathrm{b}}}{2}\sqrt{\left(\frac{k_{\mathrm{b}}}{t_{\mathrm{f}}}\left(\frac{E_{\mathrm{fT}} + E_{\mathrm{fL}}}{E_{\mathrm{fT}}E_{\mathrm{fL}}}\right)\right)} \tag{4}$$

Here E_{fl} and E_{fT} represent the elastic moduli in the axial and transverse directions of the fiber, respectively; t_{f} is the thickness of the collapsed fiber; and l and l_{b} are as previously defined.

The factor C_B determines the amount of reinforcement associated with fiber bonding. The constant k_b in the factor B (Eqn. (4)) determines the stiffness of the fiber–fiber bond. It is expected that these factors are related to the degree of fiber refining which roughens the fiber surface and causes damage or separation of the S1 layer of the cell wall. As the bond stiffness becomes very small, C_B tends to a value of 2 and the reinforcing effect vanishes. At the other extreme, the bond becomes infinitely stiff and the factor C_b tends to unity.

The bond stiffness can be measured directly as shown by Thorpe *et al.* (1976) where individual fibers are bonded to a fiber shive which is supported and the fiber is loaded by an externally applied axial load. The value of k_b measured in this way is on the order of $10^{13}\,\mathrm{Nm^{-3}}$.

The modeling of the elastic properties of the fiber depends on a number of factors related to whether or not the fibers are thin-walled and collapsed in the paper sheet and whether or not it is considered necessary to include the effects of the S_1 and S_3 layers of the cell wall. Refer to the work of Mark and Gillis (1983) and Salmen and de Ruvo (1985) for detailed information regarding the modeling of the fiber. Note that the transverse fiber modulus, E_{fT}, should be computed based on the crossing angle that the crossing fibers make relative to the primary fiber as discussed in the following.

The strain in the fiber, ε_f, is determined by the sheet strain and the degree of coupling between each fiber and the other fibers in the sheet. As shown in Fig. 1, a fiber of length λ_f is coupled to the rest of the network through the crossing fibers. The equilibrium equation for fiber mesoelement based on the equilibrium of a typical element along the fiber length is

$$\frac{d}{d\xi}\left(E_{af}\frac{d\varepsilon_f}{d\xi}\right) - k_f\varepsilon_f = -k_f\varepsilon_s \qquad (5)$$

where

$$\varepsilon_s = \varepsilon_x\cos^2\theta + 2\varepsilon_{xy}\sin\theta\cos\theta \qquad (6)$$

Equation (6) represents the normal component of strain in the sheet in the direction θ corresponding to the orientation of the fiber. The factor k_f is the effective stiffness of the coupling between the fiber of orientation θ and the other crossing fibers that couple the fiber to the network. The stiffness k_f includes the flexibility of the fibers in bending (which may be discernible for low-density sheets) and the flexibility of the fiber–fiber bond. If the idealized system of rectangular cross section straight fibers is considered,

$$k_f = \frac{2}{l\left[\dfrac{(l-l_b)^3}{12E_{fL}I_f}\left(1+\dfrac{12\mathrm{f}E_{fL}I_f}{G_fA_f(l-l_b)^2}\right)+\dfrac{2}{A_bk_b}\right]} \qquad (7)$$

where I_f represents the moment of inertia associated with the fiber cross-section, G_f represents the cell wall shear modulus, f is a factor that depends upon the fiber cross-sectional shape (it is 6/5 if the fiber is rectangular), A_b is the area of the fiber–fiber bond, and k_b is the bond stiffness.

Equations (2)–(4) and (7) can be used to estimate values of E_{af} and k_f in an idealized case. These quantities can also be considered as fundamental parameters of the system which are to be determined by mechanical property testing. In the latter case, direct measurement of paper mechanical response is used with the theoretical property relations to calculate "experimental" values of E_{af} and k_f.

It should be noted that the values of the effective axial modulus, E_{af}, and the coupling stiffness, k_f, are dependent on the orientation angle, θ, inasmuch as the network lengths, l and l_b, and the bond area, A_b, are dependent on orientation angle. In order to permit the subsequent closed-form expressions for the moduli of the sheet it is necessary to take the network lengths as their average values for the sheet. These values are dependent on the orientation distribution of the fibers as shown below.

For a prescribed sheet strain given by ε_x, ε_y, ε_{xy}, the normal component of sheet strain in the direction of the fiber, ε_s, is known. Equation (5) can be solved by assuming an arbitrary value of ε_s. The factors A_f, E_{af}, and k_f should properly be considered as varying in some random fashion. In this case, the solution to Eqn. (5) is very difficult. An alternative procedure is to select average values for A_f, E_{af}, and k_f and to determine an average fiber response. When this procedure is followed, the solution of Eqn. (5) is easily obtained as

$$\varepsilon_f = \varepsilon_s\left(1-\frac{\cosh a\xi}{\cosh aL}\right) \qquad (8)$$

$$a = \left(\frac{k_f}{A_fE_{af}}\right)^{0.5} \qquad (9)$$

L is half the fiber length and A_f is the cross-section area of the fiber.

The strain energy of a typical segment element is calculated as

$$W_e = \tfrac{1}{2}E_{af}A_f\lambda\varepsilon_s^2\eta_L \qquad (10)$$

where

$$\eta_L = 1-\frac{\tanh aL}{aL} \qquad (11)$$

and can be identified as a coupling efficiency. For strong coupling and long length, η_L approaches unity. At the other extreme of very short length or very energy of the system can be written

$$W = \int_0^x \int_0^\pi D_0 f^*_{\theta\lambda} W_\mathrm{e} \mathrm{d}\theta \mathrm{d}\lambda \tag{12}$$

where D_0 represents the number of segment elements per unit area and $f^*_{\theta\lambda} W_e \mathrm{d}\theta\mathrm{d}\lambda$ represents the probability of finding a segment having a length in the interval λ, $\lambda + \mathrm{d}\lambda$ and orientation in the interval θ, $\theta + \mathrm{d}\theta$.

Suppose that $f^*_{\theta\lambda}$ can be expressed in the form

$$\begin{aligned} f^*_\theta = {} & [f_\theta(\theta,\ a_1,\ a_2,\ a_3,\ \ldots,\ a_m)] \\ & \times [f_\lambda(\lambda_s,\ b_1,\ b_2,\ b_3,\ \ldots,\ b_m)] \end{aligned} \tag{13}$$

Here, f_θ describes the fiber orientation distribution as a function of θ depending on the m parameters, a_1, a_2, $\ldots$, a_m. Likewise, the distribution of fiber lengths, λ, depends on n parameters b_1, b_2, $\ldots$, b_n. In general, the parameters a_1, a_2, $\ldots$, a_m are assumed to be functions of λ, while the parameters b_1, b_2, $\ldots$, b_n are functions of θ.

The evidence available suggests that coupling between the two distributions is weak and therefore as an approximation, valid for at least certain papers, it can be assumed that the length and orientation distributions are independent. For this assumption, the parameters a_1, a_2, $\ldots$, a_m and b_1, b_2, $\ldots$, b_n are constants, and the strain energy per unit sheet area can be written in the form

$$W = \frac{1}{2}\frac{\omega_\mathrm{s}}{\rho_\mathrm{af}} \int_0^\pi \phi_\lambda E_\mathrm{af} \varepsilon_s^2 f_\theta \mathrm{d}\theta \tag{14}$$

where ω_s represents the basis weight, E_af represents the apparent fiber density and

$$\phi_\lambda = \int_0^\infty \eta_\mathrm{L} f_\lambda \mathrm{d}\lambda \tag{15}$$

The length of parameter ϕ_λ and the apparent fiber modulus, E_af, may depend upon the orientation direction, θ, as a result of the stresses imposed on the fibers during drying of the sheet. Analytically, it can be assumed that the apparent fiber modulus, E_af, depends upon the shrinkage and restraint conditions present during drying in the following way:

$$E_\mathrm{af} = E_{\mathrm{af}_0}(1 + H\varepsilon_\mathrm{ND}) \tag{16}$$

where E_{af_0}, a constant, represents the apparent fiber modulus in the absence of drying restraint, H is a constant that predicts the magnitude of stiffening, and

$$\varepsilon_\mathrm{ND} = (\alpha_x M + \varepsilon_{x\mathrm{D}})\cos^2\theta + (\alpha_y M + \varepsilon_{y\mathrm{D}})\sin^2\theta \tag{17}$$

Here $\alpha_x M$ and $\alpha_y M$ represent the sheet shrinkage strains during unrestrained drying in the x and y directions, and $\varepsilon_{x\mathrm{D}}$ and $\varepsilon_{y\mathrm{D}}$ represent the sheet strains applied or allowed during drying. Thus, for unrestrained shrinkage conditions, $\varepsilon_{x\mathrm{D}} = -\alpha_x M$ and $\varepsilon_{y\mathrm{D}} = -\alpha_y M$. It is evident that ε_ND is related to the magnitude of restraint that a fiber of orientation θ would be subjected to during the drying of the sheet. In the following analysis, it is further assumed that ε_ND is positive, that is, the fibers are loaded in tension as a result of any drying restraint.

Experimental results of the effects of drying restraint were performed by Setterholm and Chilson (1965). Data taken from their work (cf. Perkins and Mark (1981)) indicates that the parameter H has values that lie in the range of 10 to 30. Experimental evidence of the effect of drying restraint on the properties of the fibers in the network was shown by Wuu *et al.* (1991).

With the help of Eqn. (16), integration of Eqn. (14) can be easily carried out, if the length factor, ϕ_λ, is independent of θ. In fact, ϕ_λ does depend on θ, because the coupling parameter a that appears in η_L depends on E_af. If one assumes that the influence of variation in E_af through θ in ϕ_λ has an insignificant effect on the prediction of the elastic moduli of the sheet, there is no serious analytical error committed in assuming the ϕ_λ is independent of θ.

The stress–strain relations for the sheet can be obtained from the strain energy function W after carrying out the integration through the relations

$$\tau_x = \frac{\partial W}{\partial \varepsilon_x} \qquad \tau_y = \frac{\partial W}{\partial \varepsilon_y} \qquad \tau_{xy} = \frac{1}{2}\frac{\partial W}{\partial\ \varepsilon_{xy}} \tag{18}$$

where τ_x, τ_y, τ_{xy} represent the force per unit edge length of the paper sheet.

To express conveniently the resulting elastic moduli it is desirable to express the orientation distribution $f_\theta(\theta)$ in terms of its Fourier series expansion. When the x direction is one of the axes of elastic symmetry, for example the machine direction, and $\theta = 0°$ corresponds to this direction,

$$\begin{aligned} f_\theta(\theta) = {} & \frac{1}{\pi}[1 + a_1 cos2\theta + a_2 cos4\theta \\ & + a_3 cos6\theta + \cdots + a_n cos2n\theta + \cdots] \end{aligned} \tag{19}$$

After substitution of Eqns. (19) and (16) in Eqn. (14) and subsequently using Eqn. (18), the Young's moduli E_x^*, E_y^*, the shear modulus G_{xy}^*, and the Poisson ratios $\upsilon_{xy,}\upsilon_{yx}$ corresponding to plain stress loading of the sheet are found to be

$$\begin{aligned} E_x^* = {} & \frac{1}{16}\frac{\omega_\mathrm{s}}{\rho_\mathrm{af}}\phi_\lambda E_{\mathrm{af}_0}\Big[(1 + \langle e \rangle)(6 + 4a_1 + a_2) \\ & + \frac{e_\Delta}{2}(8 + 7a_1 + 4a_2 + a_3)\Big]\left[1 - v_{xy}v_{xy}\right] \end{aligned} \tag{20a}$$

$$E_y^* = \frac{1}{16}\frac{\omega_{\mathrm{s}}}{\rho_{\mathrm{af}}}\phi_\lambda E_{\mathrm{af}_0}\Big[(1+\langle e\rangle)(6-4a_1+a_2) + \frac{e_\Delta}{2}(8-7a_1+4a_2-a_3)\Big][1-\nu_{xy}\nu_{xy}] \tag{20b}$$

$$\nu_{xy} = \frac{(1+\langle e\rangle)(2-a_2)+\dfrac{e_\Delta}{4}(a_1-a_3)}{(1+\langle e\rangle)(6-4a_1+a_2)-\dfrac{e_\Delta}{2}(8-7a_1+4a_2-a_3)} \tag{20c}$$

$$\nu_{yx} = \frac{(1+\langle e\rangle)(2-a_2)+\dfrac{e_\Delta}{4}(a_1-a_3)}{(1+\langle e\rangle)(6+4a_1+a_2)+\dfrac{e_\Delta}{2}(8+7a_1+4a_2+a_3)} \tag{20d}$$

$$G_{xy}^* = \frac{1\,\omega_s}{16\rho_{\mathrm{af}}}\phi\lambda E_{\mathrm{af}_0} \times\Big[(1+\langle e\rangle)(2-a_2)+\frac{e_\Delta}{2}(a_1-a_3)\Big] \tag{20e}$$

where

$$\langle e\rangle = \tfrac{1}{2}H[(\alpha_x M + \varepsilon_{x\mathrm{D}}) + (\alpha_y M + \varepsilon_{y\mathrm{D}})] \tag{21a}$$

$$e_\Delta = \tfrac{1}{2}H[(\alpha_x M + \varepsilon_{x\mathrm{D}}) - (\alpha_y M + \varepsilon_{yD})] \tag{21b}$$

An inspection of Eqns. (20a–e) indicates that the elastic constants depend only on the first three parameters, a_1, a_2, a_3, that occur in the Fourier series expansion for f_θ. Furthermore, if the influence of drying restraint is not incorporated or if the sheet is dried under conditions of no restraint, the elastic moduli depend only on the parameters a_1 and a_2. For this condition, Eqns. (20a–e) are essentially the same as those given by Cox (1952) when $\theta = 0^\circ$ corresponds to a direction of elastic symmetry in the sheet. (Note, however, that the effects of fiber–fiber bonding and the effect of finite fiber length are taken into account.)

The number of coefficients of the Fourier series expansion for f_θ that are necessary for predicting the elastic properties, therefore, depends on the functional relation between the apparent fiber modulus, E_{af}, and orientation, which in turn depends on the procedures and conditions of drying. Incorporation of any other phenomena affecting the angular depencence of the strain energy stored in the sheet may change the number of Fourier coefficients appearing in the elastic constants.

It is theoretically possible to describe any fiber orientation distribution in terms of the parameters a_1, a_2, a_3, … of the Fourier series expansion. However, a single-parameter function would simplify the expressions for the elastic moduli. The reduction to one parameter can be accomplished very simply by truncating the Fourier series and retaining the single parameter a_1. However, reduction to a coefficient in the Fourier series expansion can have some significant effects. For example, for freely dried paper, the shear modulus should decrease with increased anisotropy, but this reduction effect will be predicted only if the coefficient a_2 is retained in the expansion of Eqn. (19). A single-parameter function would predict the shear modulus to be independent of the degree of anisotropy of the paper.

An alternative approach is to employ a distribution function that has only one shape parameter. For a general discussion of fiber orientation distribution refer to Perkins and Mark (1981). The two most suitable single-parameter distributions are the elliptical and the von Mises distributions. The distributions and their corresponding values of a_1, a_2, a_3 are given in Table 1.

2. *Application of Theory to Prediction of Elastic Moduli of a Paper Sheet*

Equations (20a–e) can be used to estimate the effects of changes in structural parameters on the elastic properties of paper, or to estimate from experimental measurements the parameters of the system. As an

Table 1
Parameters for the elliptical and von Mises fiber orientation distributions.

	Parameter		
	a_1	a_2	a_3
Elliptical $f_\theta = \dfrac{\zeta}{\pi}\left(\dfrac{1}{\cos^2\theta+\zeta\sin^2\theta}\right)$	$2\left(\dfrac{\zeta-1}{\zeta+1}\right)$	$2\left(\dfrac{\zeta-1}{\zeta+1}\right)^2$	$2\left(\dfrac{\zeta-1}{\zeta+1}\right)^3$
von Mises[a] $f_\theta = \dfrac{1}{\pi I_0(\kappa)}\mathrm{e}^{\kappa\cos2\theta}$	$2\dfrac{I_1(\kappa)}{I_0(\kappa)}$	$2\dfrac{I_2(\kappa)}{I_0(\kappa)}$	$2\dfrac{I_3(\kappa)}{I_0(\kappa)}$

[a] $I_n(\kappa)$ = modified Bessel function of the first kind and order n.

example, if the values of E_x and E_y can be obtained from measurements of the Young's moduli in the machine and cross-machine directions, the remaining in-plane elastic moduli can be estimated.

To use Eqns. (20a–e) in this example, it is necessary to investigate the influence of drying restraint. The stiffening parameter H of Eqn. (21) has been estimated by experiment (Setterholm and Chilson 1965) to lie in the range 10–30. From a knowledge of the drying restraint conditions and sheet shrinkage behavior, the values of $\langle e \rangle$ and $\langle e_\Delta \rangle$ can be calculated. The ratio of the experimental Young's moduli for the machine and cross-machine directions can be equated to the ratio E_x^*/E_y^* from Eqns. (20a,b). The elliptical or von Mises distribution can be written in a Fourier series form (Table 1). The first three coefficients of the expansion can then be taken as a_1, a_2, a_3, respectively. The inversion of this relation would yield a value of the concentration parameters for orientation ζ or κ can then be used to obtain a_1, a_2, a_3 and subsequently the values of ν_{xy}, ν_{yx}, and G_{xy}. The measurement of two Young's moduli then permits all the in-plane elastic moduli for the sheet to be estimated. The calculation is best performed for several different assumptions regarding the variable in Eqn. (21), to ascertain the reliability of the assumptions.

The choice of fiber orientation distribution function can have significant effects on predictions obtained from Eqns. (20a–e). For a given fiber orientation, predictions of ν_{xy}, ν_{yx}, and G_{xy} from experimental measurements of E_x and E_y may be different for the elliptical distribution from those for the von Mises distribution. While the elliptical distribution or the single-parameter cosine distribution is the easiest to apply, the von Mises distribution is considered to provide more accurate estimates (Perkins and Mark 1981).

In the case of fiber segment length distributions, the Erlang distribution has been found to be quite satisfactory. This, the function f_{λ_s} can be taken as

$$f_{\lambda_s} = \frac{(\lambda_s/b)^{c-1}\exp(-\lambda_s/b)}{b[(c-1)!]} \tag{22}$$

Here, b and c are determined by fitting the distribution to the experimental data. The product bc is equal to the mean value of λ_s. The standard deviation for the distribution is $b\sqrt{c}$. For further information, refer to Perkins and Mark (1981).

The distance between fiber crossings of fiber-to-fiber bonds along a given fiber, $l(\theta)$, is an important property of the fiber network. Assuming that the fibers are all straight, the bond centroid distance depends on the direction of the fiber and the overall orientation distribution of the fibers in the network. The bond centroid distance is determined by the number of fibers that cross a given fiber. This problem was studied by Kallmes and Corte (1960) and by Komori and Makishima (1977) who showed that

$$n(\theta) = \frac{2D_f N\lambda^2}{V} J(\theta) \tag{23}$$

where

$$J(\theta) = \int_0^\pi f(\theta)|\sin(\theta - \alpha)|\mathrm{d}\alpha \tag{24}$$

Here, D_f denotes the fiber diameter, or in the case of a flattened fiber, the *z*-direction thickness of the fiber. N represents the number of fibers in a volume V. The function $J(\theta)$ represents the average value of the sine of the crossing angle, $\theta-\alpha$, for a fiber of orientation θ. The function $J(\theta)$ can be expressed in series form as shown by Chang (1983)

$$J(\theta) = \frac{2}{\pi}\left[1 - \sum_{n=1}^{\infty} \frac{a_n}{(2n-1)(2n+1)}\cos(2n\theta)\right] \tag{25}$$

The rationale for Eqn. (23) follows from the observation that $2D_f\lambda(\lambda\sin\beta)$ represents the volume occupied by the two crossing fibers times the projected length of the crossing fiber. Thus, the average number of contacts is the number of fibers, N, in a volume, V, times the ratio of the volume associated with one crossing to the total volume of fiber material.

The average distance between bond centroids, $l(\theta)$, along a fiber of orientation θ is taken as the reciprocal of the average number of crossings per unit length. Therefore,

$$l(\theta) = \frac{V}{2D_f N\lambda J(\theta)} \tag{26}$$

The average number of crossings considered over the set of all fibers is

$$\bar{n} = \frac{2D_f N\lambda^2}{V} I \tag{27}$$

where

$$I = \int_0^\pi f(\alpha)J(\alpha)\mathrm{d}\alpha \tag{28}$$

or in series form as shown by Chang (1983)

$$I = \frac{2}{\pi}\left[1 - \sum_{n=1}^{\infty} \frac{a_n}{(2n-1)(2n+1)}\right] \tag{29}$$

The overall average distance between bond centroids for the network is

$$l = \frac{V}{2D_f N\lambda I} \tag{30}$$

The quantity $N\lambda$ represents the total length of fiber in a volume, V. If $M = \rho_f A_f N\lambda$ represents the mass in volume V, then the expressions for bond centroid distance can be written,

$$l(\theta) = \frac{\rho_{\rm f}}{2\rho_{\rm s}} \frac{A_{\rm f}}{D_{\rm f}} \frac{1}{J(\theta)} \tag{31}$$

$$l = \frac{\rho_{\rm f}}{2\rho_{\rm s}} \frac{A_{\rm f} 1}{D_{\rm f} I} \tag{32}$$

When the fibers are assumed to be completely flattened, then $D_{\rm f}$ is the thickness $t_{\rm f}$ and $A_{\rm f} = w_{\rm f} t_{\rm f}$ and the bond centroid distances can be written

$$l(\theta) = w_{\rm f} \frac{\rho_{\rm f}}{2\rho_{\rm s}} \frac{1}{J(\theta)} \tag{33}$$

$$l = w_f \frac{\rho_f}{2\rho_s} \frac{1}{I} \tag{34}$$

The fiber crossing angle can be obtained from the definition of $J(\theta)$ since this quantity represents the average value of the sine of the angle between two crossing fibers for a fiber of given orientation angle θ relative to the reference direction. Therefore, the average crossing angle β is given by

$$\beta = \arcsin(J(\theta)) \tag{35}$$

The bond area is calculated from the crossing angle. Therefore,

$$A_{\rm b}(\theta) = \frac{w_{\rm f}^2}{J(\theta)} \tag{36}$$

and the average bond area for fibers of all orientations is

$$A_{\rm b} = \frac{w_{\rm f}^2}{I} \tag{37}$$

These network properties are to be used in the calculations for the fiber coupling parameter $k_{\rm f}$ and the effective fiber modulus $E_{\rm af}$ in Eqns. (2) and (7), respectively.

When the relations provided in Eqn. (20) are to be used to predict the sheet elastic moduli, the values of l, and $A_{\rm b}$ are to be taken from Eqns. (34) and (37) because these relations relate the average properties of the sheet over all directions. The portion of the fiber that is bonded, $l_{\rm b}$, is defined using Eqns. (36) and (37) as the apparent length necessary in order that the product $l_{\rm b} w_{\rm f}$ is equal to the bond area. On the other hand, the methods leading to Eqn. (20) can be performed numerically with the angle dependence of bond centroid distance, bond length, and bond area taken into account using Eqns. (33) and (36) in the expressions for $E_{\rm af}$ and $k_{\rm f}$. The numerical approach makes it possible to avoid the simplifying assumption of independence of angular dependent and fiber length effects that are discussed above.

Using the relations above, it can be shown that when the sheet density, $\rho_{\rm s}$, is equal to one-half the fiber density, $\rho_{\rm f}$, the fibers will be in contact with adjacent fibers on one side of the fiber or the other and the unbonded length goes to zero. Given that the fiber density is 1500 kg m^{-3}, the network relations above predict that fibers will have no unbonded or free length when the sheet density is above 750 kg m^{-3}. For paper sheets characterized by a density in excess of this figure, the mesoelement model depicted by Fig. 1 should be modified so that there is no free fiber length. This would suggest that the mesoelement should be taken as the fiber itself. Therefore, for sheet density in excess of 750 kg m^{-3} the relations for the elastic moduli (Eqn. (20)) can still be used as long as the effective axial fiber modulus from Eqn. (2) is taken as the axial fiber modulus, $E_{\rm fL}$, and the network coupling stiffness, $k_{\rm f}$, is taken as

$$k_{\rm f} = \frac{A_{\rm b} k_{\rm b}}{l} \tag{38}$$

When the procedure described above is used for predicting the elastic moduli of sheets with sheet density in excess of 750 kg m^{-3} it is clear that the prediction neglects the transverse and shearing strains in the fiber. The laminate approach was explored by Schulgasser and Page (1988) and by Page and Schulgasser (1989) in order to take into account the transverse and shearing contribution. They showed that the transverse properties of fibers are especially important as regards the prediction of the in-plane shear modulus and the Poisson ratios of the sheet. In the works of Lu *et al.* (1995), Lu and Carlsson (1996), and Lu *et al.* (1996) the authors develop what they refer to as a mosaic model that incorporates the in-plane axial and transverse elastic moduli of the fiber. This approach makes use of the finite element method to analyze a representative element that is characteristic of the fiber segments and bonds that comprise the paper network. A model for moderate to high density paper was also proposed by Uesaka (1990) where the transverse strain effects were taken into account by modifying the theory leading to Eqns. (20) to incorporate transverse fiber strain in a similar fashion to the modeling of the axial strain. This work takes into account the contribution of the fiber-to-fiber bonds, but neglects the fiber Poisson and shearing effects. An approximate model that accounts for transverse strain, Poisson and shearing effects, and the fiber–fiber bonding has been proposed by Perkins (1999).

3. *Nonlinear Geometry and Constitutive Effects*

The assumption of linear elastic behavior of the fiber and bond materials can be relaxed while using essentially the same approach as that outlined above. The mechanical response of paper when loaded uniaxially, as in the example of a simple tensile test, was studied by Ramasubramanian and Perkins (1988) who developed algebraic expressions for predicting the uniaxial stress–strain response. The expressions are sufficiently complex that their evaluation by computer is necessary. This work was extended to account for moderate to high density papers by Perkins *et al.* (1991). Subsequently, Sinha and Perkins (1995) adapted the mesoelement approach to provide a constitutive model for use in finite element analysis.

The basic approach outlined above can be modified to take into account a number of nonlinear effects if one does not require that closed form expressions be obtained for the prediction of mechanical response. In other words, the mesoelement approach can be extended by performing the calculations numerically rather than in algebraic form. The effects of fiber curl, fiber buckling, nonelastic constitutive models for the bond and fiber components, as well as certain aspects of failure at the mesoelement level can be studied when the mesoelement model is analyzed by the finite element method. The basic approach has been presented by Perkins (1999).

See also: Paper: Porosity#; Paper, Wetting and Penetration of Liquids into#; Paper: Strength and Stiffness*

Bibliography

Algar W H 1996 In: Bolam F (ed.) *Consolidation of the Paper Web*. Technical Section of the British Paper and Board Makers' Association, Cambridge, Vol. 2, pp. 814–49

Chang J 1983 Syracuse University, Syracuse, NY

Deng M, Dodson C T J 1994 *Paper: An Engineered Stochastic Structure*. Tappi Press

Dodson C T J 1973 In: Bolam F (ed.) *The Fundamental Properties of Paper Related to its Uses*. Technical Association of the British Paper and Board Makers' Association, Cambridge, Vol. 1, pp. 202–26

Kallmes O 1972 In: Jayne B A (ed.) *Theory of Design of Wood and Fiber Composite Materials*. Syracuse University Press, Syracuse, NY, pp. 157–76

Kallmes O 1972b In: Jayne B A (ed.) *Theory of Design of Wood and Fiber Composite Materials*. Syracuse University Press, Syracuse, NY, pp. 177–96

Kallmes O, Corte H 1960 The structure of paper, I. *Tappi*. **43**, 737–52

Komori T, Makishima K 1977 *Textile Res. J.* **47**, 13–7

Lu W, Carlsson L A 1996 Micromodel of paper, II. *Tappi J.* **79**, 205–10

Lu W, Carlsson L A, Andersson Y 1995 Micromodel of paper. *Tappi J.* **78**, 155–64

Lu W, Carlsson L A, de Ruvo A 1996 Micromodel of paper, I. *Tappi J.* **79**, 197–203

Mark R E, Gillis P P 1983 In: Mark R E (ed.) *Handbook of Physical and Mechanical Testing of Paper and Paperboard*. Marcel Dekker, New York, Vol. 1, pp. 409–95

Niskanen K (ed.) 1998 *Paper Physics*. Fapet Oy, Hensinki, Finland

Page D H, Schulgasser K 1989 In: Perkins R W (ed.) *Mechanics of Cellulosic and Polymeric Materials*. ASME, New York, Vol. AMD-Vol. 99, MD-Vol. 13, pp. 35–9

Perkins R 1990 In: Caulfield D F, Passaretti J D, Sobczynski S F (eds.) *Materials Interactions Relevant to the Pulp, Paper, and Wood Industries*. Materials Research Society, San Francisco, CA, Vol. 197, pp. 99–118

Perkins R W 1983 In: Mark R E (ed.) *Handbook of Physical and Mechanical Testing of Paper and Paperboard*. Marcel Dekker, New York, Vol. 1, pp. 23–5

Perkins R W 1999 In: *ASME Summer Mechanics Conference-Symposium on Mechanics of Cellulosic Materials*, Vol. AMD-Vol. 231

Perkins R W (ed.) *American Society of Mechanical Engineers*, Blacksburg, VA, MD-Vol. 85

Perkins R W, Mark R E 1981 The role of fundamental research in papermaking. In: Brander J (ed.) *Trans. Symp.* Cambridge: September *1981*, Mechanical Engineering Publications Limited, London, Cambridge, UK, Vol. 1, 479–525

Perkins R W, Sinha S, Mark R E 1991 In: Mark R E (ed.) *Int. Paper Physics Conf.* TAPPI, Kona, Hawaii, Vol. 2, pp. 413–35

Qi D 1997 *Tappi J.* **80**, 165–71

Ramasubramanian M K, Perkins R W 1988 *J. Eng. Mater. Technol.* **110**, 117–23

Ramasubramanian M K, Wang Y Y 1999 In: Perkins R W (ed.) *Mechanics of Cellulosic Materials*. ASME, New York, pp. 31–42

Salmen L, de Ruvo A 1985 *Wood Fiber Sci.* **17**, 336–50

Schulgasser K, Page D H 1988 *Composites Sci. Technol.* **32**, 279–92

Setterholm V C, Chilson W A 1965 *Tappi J.* **48**, 634–40

Sinha S S, Perkins R W 1995 In: Perkins R W (ed.) *Mechanics of Cellulosic Materials*. ASME, New York, Vol. MD-Vol 60, AMD-Vol 209

Thorpe J, Mark R E, Eusufzai A R K, Perkins R W 1976 *Tappi J.* **59**, 96–100

Uesaka T 1990 In: Perkins R W (ed.) *Mechanics of Wood and Paper Materials*. ASME, New York, Vol. AMD-Vol. 112, MD-Vol. 23, pp. 29–35

Van den Akker J A 1970 *Tappi J.* **53**, 388–400

Van den Akker J A 1972 In: Jayne B A (ed.) *Theory of Design of Wood and Fiber Composite Materials*. Syracuse University Press, Syracuse, NY, pp. 197–218

Wuu F, Mark R E, Perkins R W 1991 In: Mark R E (ed.) *International Paper Physics Conference*. Tappi, Kona, Hawaii, Vol. 2, pp. 663–71

R. Perkins
Syracuse University, New York, USA

Papermaking

The operations involved in converting fibers (see *Paper: Pulping and Bleaching**; *Pulp and Paper: Nonwood Sources*#) and other materials (see *Pulp and Paper: Nonfibrous Components**) into a sheet of paper or paperboard are collectively known as papermaking. Paper having a width up to 10 m is manufactured on a continuous basis on paper machines as long as 150 m. Production speed is 33 m s^{-1} for some grades.

Papermaking has evolved from an art practiced by hand to a complex and technical manufacturing process with a high degree of automation (see *Paper: History of Development*[#]).

The principal steps of papermaking include stock preparation, forming to create a uniform wet web, pressing to remove liquid water, drying to evaporate most of the remaining water and consolidate the product, calendering to reduce the sheet thickness and improve the surface smoothness, and reeling to collect the dry product. All these operations are carried out on a single, large paper machine. Additional operations, such as winding to produce smaller rolls of paper suitable for printing or for further converting, the application of various additives onto the sheet, or supercalendering, may be carried out on- or off-machine to produce different grades of paper (see *Paper Products: Classification**).

Many of the papermaking operations involve water removal. The development of sheet dryness as it proceeds along a typical paper machine is shown in Fig. 1. In the forming section, about 95% of the water ejected from the headbox is removed as the web solids content is rapidly increased from less than 1% to a value ranging anywhere from 13% to 25%. This is followed by pressing, which typically achieves a web solids content somewhere between 38% and 50%. Drying is used only for the last stage of water removal, resulting in a final moisture content of the paper between 2% and 10%.

Paper is truly an engineered material that is manufactured according to the requirements of its final use. Many properties, such as surface texture, color, strength, stretch, absorbency, permeability, translucency, grease or moisture resistance, electrical conductivity, permanency, and stiffness, can be imparted by carefully tailoring the manufacturing process.

1. Forming

Forming is the critical step of the papermaking process. The objective of forming is to drain water from a dilute fiber suspension deposited on a forming fabric, and to create a uniform fiber mat that has a sufficient consistency to be passed to the press section of the paper machine. Forming has a strong influence on a range of paper properties. A sheet with ideal formation would have a uniform mass distribution in the plane of the sheet with no flocs and no difference in fiber orientation at all points across the machine. Forming sets the fiber orientation in the sheet, which

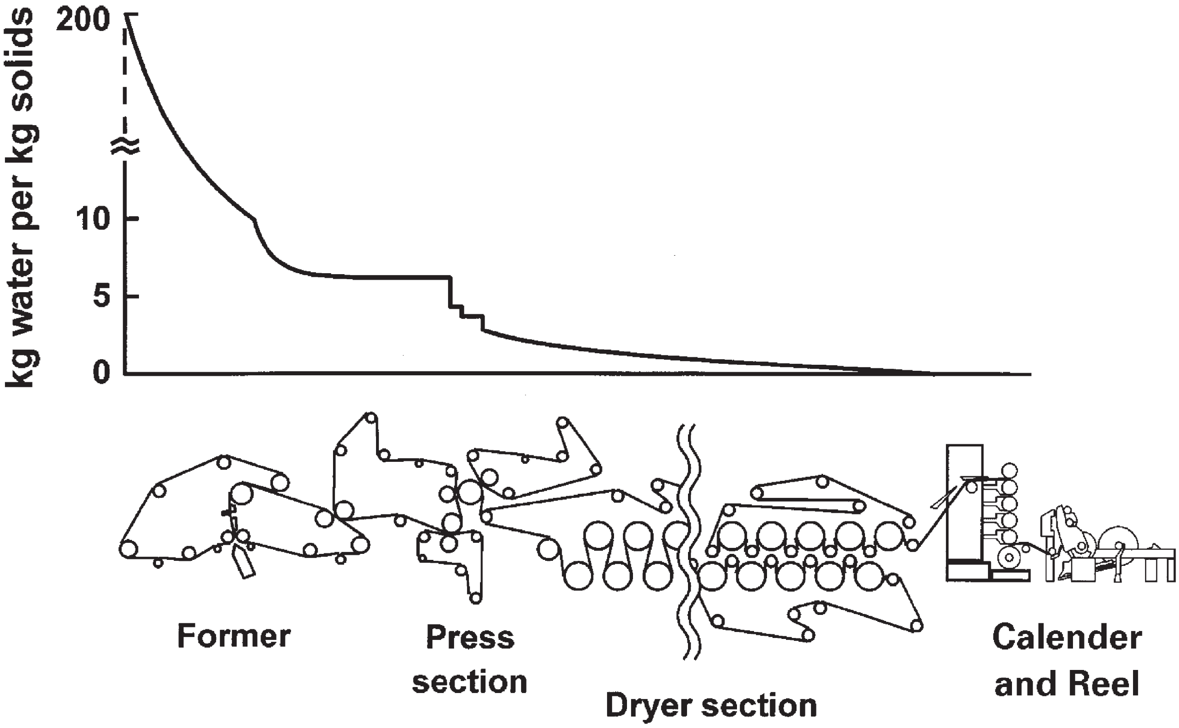


Figure 1
A large amount of dewatering occurs as the paper web proceeds down a typical paper machine, which includes forming, pressing, drying, calendering, and reeling sections.

will influence the directionality of paper properties. As well, forming determines the fines and filler distribution in the *z* direction, which determines the two-sidedness and surface smoothness of the product. The characteristics of the web created in the forming process can be linked to many of the final paper properties (see *Paper: Strength and Stiffness**). It is thus useful to consider not only the process and equipment but also how the forming and drainage processes are coupled to quality.

Stock preparation is a series of operations in which the paper machine furnish is prepared for papermaking (see *Paper: Pulping and Bleaching**). The operations involved might include dispersion of dry pulp or paper in water, refining of pulp, blending of several pulps and the addition of fillers, dyes, or other additives (see *Pulp and Paper: Nonfibrous Components**). The prepared stock at a consistency of about 3% is stoved in a machine chest and is referred to as "thick stock."

1.1 Approach Flow System

The approach flow system includes all pumps, pipes, and apparatus through which the stock flows from the machine chest to the headbox. The function of the approach flow system is to provide a continuous, uniform flow of cleaned, screened, and deaerated fibrous suspension to the paper machine. Screens and centrifugal cleaners in the approach flow system remove large and heavy contaminants from the pulp. Some additives such as retention aids might be added to the pulp in the flow approach system. Dissolved and dispersed air may interfere with the formation, and is often removed from the pulp suspension. In the approach flow system, water drained during forming is used to dilute the thick stock from the machine chest from a mass concentration, or consistency, of about 3% to 1%. This low consistency is required to minimize fiber flocculation, which impairs the quality of formation.

1.2 Distributor and Headbox

The stock arrives through a pipe at the side of the paper machine. The function of the distributor and headbox is to convert the flow from the cross-direction to the machine direction and to spread the pipeline flow into a jet of full machine width, but only a few centimeters thick. The jet is then deposited onto the forming fabric. The uniformity of the jet has a considerable effect on the quality of paper formation. It is critical that the thickness and velocity of the jet is uniform across its full width, since this determines the uniformity of product basis weight.

Tapered headers are the most common type of flow distributor. The flow turns from the incoming pipeline into a bank of tubes, or holes in a drilled plate. The taper provides a uniform pressure across the entries to the tube bank, so that the flow will spread equally across the paper machine. The recirculation flow at the narrow end of the manifold is adjusted to ensure a uniform pressure across the machine.

The stock flows through holes or tubes directly into the headbox, which may be either an air pad or hydraulic headbox. The air pad headbox was dominant until the early 1970s. In this design the flow from the tubes emerges into a chamber equipped with one or several rotating, perforated rolls. These rolls disperse flocs of fibers and provide a flow resistance that blends the jets emerging from individual tubes and minimizes side flows within the headbox, which could cause nonuniformity in the jet flow. This dampens out pressure disturbances in the incoming flow, which might otherwise produce basis weight variations. This headbox design is inappropriate for modern, wide, high-speed paper machines because its large chamber is not compatible with high stock pressures, and the long perforated rolls are easily deformed.

Hydraulic headboxes, commonly installed on modern paper machines, have no perforated rolls and usually no air pads. Instead, the flow from the distributor proceeds through deflocculating elements into a converging section, where it is accelerated to a speed close to that of the paper machine. Since the hydraulic headbox does not have a large-volume chamber, it can withstand the high stock pressures required to accelerate the jet to a high machine speed.

The jet of fiber suspension leaves the headbox through an adjustable slice consisting of two thin metal plates, referred to as the upper and lower lips. The basis weight, or amount of fiber mass deposited per unit area, is proportional to the flow of thick stock from the machine chest. The consistency of the suspension is controlled by the slice opening. Large adjustments to the slice lip opening are made when switching between different paper grades. Smaller adjustments are used to provide a uniform basis weight profile in the cross-machine direction. These adjustments are made at discrete locations, typically every 10–15 cm across the paper machine width. However, a local correction at one slice position can lead to secondary effects in the flow of adjacent control zones, resulting in nonuniformities in the paper basis weight and fiber orientation. Very uniform basis weight profiles can be accomplished only with the assistance of sophisticated computer control.

The uniformity of basis weight profiles can also be adjusted by adding varying amounts of fresh water or white water entering the headbox, to modifiy the local consistency rather than the flow rate. Dilution control thus provides basis weight uniformity without affecting the local flow patterns. The positioning of the slice lips are also used to set the discharge angle of the jet with respect to forming fabric. This angle affects formation, porosity, and other properties.

1.3 Formers

There are four main classes of former: cylinder-vat-formers, fourdriniers, top formers, and gap formers.

(*a*) *Cylinder-vatformers.* Cylinder vatformers were common in the past, but they are not suitable for high speed operations. Today they are used mainly on multiply board machines.

(*b*) *Fourdrinier formers.* The fourdrinier was dominant from the early 1800s until the late 1960s. In the classic fourdrinier design, the jet from the headbox impinges on the forming fabric at the forming board. Substantial dewatering occurs due to the momentum of the jet. Small differences in the speed of the jet and the fabric speed (termed "rush" or "drag") will induce a machine direction alignment in the fibers. Consequently, the tensile strength will be much greater in the machine direction of the paper than in the cross-machine direction.

The forming fabric carries the suspension over foils, which generate pulses of suction and promote drainage of the water. The leading edge of successive foils skims water from the fabric surface. The pulses also cause lateral movement of the fiber suspension, which promotes even distribution of fiber and improved formation. In the past, table rolls were used to support the forming fabric. However, at high machine speeds table rolls generate excessive pressure pulses, which can disrupt the fibrous mat. Table rolls can still be found on some old, slow machines.

The resistance of the mat to drainage increases as water is drawn off. The mat is then drained over suction boxes with progressively higher levels of vacuum to a consistency of 8–10%, and over a suction couch to reach a solids content of 15–25%. The design of the forming fabric, the size of the individual filaments, their weave, and the number of layers in the forming fabric will affect the drainage rates, retention of fines and fiber, the surface quality of the paper, and other paper properties. Fabrics must be carefully selected for each application.

A significant problem with the fourdrinier is that drainage occurs through only one side of the fiber mat. This side might be depleted of fines, filler, and other small particles, which can lead to problems with print quality, curl, and other properties.

(*c*) *Hybrid or top formers.* The top former or "hybrid" former reduces the two-sidedness of paper, and increases the dewatering capacity of a paper machine by draining water from both sides of the mat. This style of former consists of a conventional fourdrinier table having a second fabric applied on top of the mat at the point where consistency of the web is about 2%. The fibrous mat sandwiched between the lower and upper forming fabrics proceeds around a curved forming roll and/or series of forming blades to produce positive drainage pressures. The pressures can be applied to either side so that an appropriate program of pressures minimizes two-sidedness.

Top formers were installed on many conventional fourdriniers in the early 1980s. However, they did not completely eliminate two-sidedness, since the initial drainage still occurs from only one side.

(*d*) *Gap or twin wire formers.* The gap former uses two fabrics for the full length of the forming section to minimize two-sidedness and maximize drainage. It is the dominant choice for new formers where surface quality and machine speed is critical.

A typical configuration has a hydraulic headbox with the jet impinging at the nip where two fabrics come together. The previous discussion of fabric selection, jet impingement, and jet to fabric speed differentials apply in a similar manner to the fourdrinier. However for a gap former, the fiber mat is contained between the two forming fabrics that wrap around a forming roll and a series of forming blades, mounted on a forming shoe, before going over suction boxes and the couch roll. The tension of the fabrics and the curvature of the forming roll and blade shoe create a positive dewatering pressure. Associated pressure gradients provide an opportunity to redistribute fibers and improve formation. The quality of the paper sheet is strongly dependent on the quality of pulp, the uniformity of the headbox jet, and the design of forming fabrics. While roll–blade formers are most common on new machines, some gap formers have only blades or only a pair of rolls.

2. Pressing

The key objective of pressing is water removal from the wet web, although two other results of pressing, namely improvements of wet web strength and product quality, are of equal importance. On a modern paper machine, the web leaving the former is weak and readily damaged or broken as it proceeds at high velocity through the subsequent machine parts. In the press section, the solids content of the sheet is increased to about 43% on machines using roll presses, and up to 50% on machines using at least one shoe press. The sheet compaction and water removal that occur in pressing cause a three- to fourfold increase in wet sheet tensile strength.

Pressing also has an important influence on the finished product properties. The network of wet fibers leaving the former is bulky and contains a large volume of air. The product obtained by drying an unpressed, is deficient in wet web has a low density, strength, and smoothness, and is poorly suited for use as a printing base or packaging product. Only some hygienic tissue papers are made from unpressed webs.

Upon pressing the wet web, malleable wet fibers are forced into close proximity, air and most of the water separating the fibers are expelled, tubular fibers are flattened, and the fiber-to-fiber contact area is greatly increased. Paper obtained by drying a pressed web is much denser, stronger, and smoother than that obtained from unpressed webs.

2.1 *Roll and Shoe Presses*

The press section of a paper machine generally consists of two to four press nips, each consisting of two counter-rotating hydraulically loaded press rolls. The wet sheet is compressed in each press nip at linear loads ranging from $30\,\mathrm{kNm^{-1}}$ to $200\,\mathrm{kN\,m^{-1}}$. The sheet is usually pressed along with a permeable press fabric to provide a means for removing the water expelled from the sheet. The press fabric is an endless loop that continuously passes through the press nip. In single-felted presses, one side of the sheet is in a direct contact with a press roll, while the other contacts the press fabric. The press roll backing the press fabric is usually vented, that is, equipped with holes or grooves to receive the water escaping from the press fabric. Some press rolls have a shell made permeable by drilled holes and a suction box located inside the roll. Elastic covers made of hard rubber or polyurethane often cover the vented rolls. The press roll contacting the wet web must be smooth, since any roughness of the roll will be replicated on the paper surface.

The press nip formed by a roll press is about 20–40 mm wide, resulting in a nip residence time of only 1 m s at a machine speed of $30\,\mathrm{m\,s^{-1}}$. If the nip residence time is too short for adequate water removal the moisture content of the sheet leaving the press may be too high. Excessive hydraulic pressure generated in a nip under these conditions may damage or break the sheet by so-called ‘‘sheet crushing.’’ A sheet with a higher solids content can withstand higher pressure and the press nip load is thus increased from the first to the last press.

Two strategies are used to improve water removal and to prevent sheet crushing, especially on machines with very short nip residence times such as those machines operating at high speeds or those producing heavy grades of paper and board which require a lot of dewatering. Press nips of such machines may be double-felted, which means that the sheet is pressed between two permeable press fabrics, allowing water to leave the sheet through both sides. The second strategy is the replacement of one of two press rolls with a hydraulically loaded shoe. The shoe is shaped in such a way that it matches the contour of the press roll. The nip formed in this manner can be about 300 mm wide, and the time available for sheet compression and water removal is about 10 times greater than in nips formed by conventional roll presses.

2.2 *Press Section Configurations*

Sheet breaks are a common problem that limits the speed and productivity of many machines. Most frequently, the sheet break occurs when the weak wet web is for the first time transferred unsupported between two machine parts in a so-called ‘‘open draw.’’ Until the 1950s the first open draw occurred between the former and the first press, and frequent breaks at this point limited the machine speed to about $11\,\mathrm{m\,s^{-1}}$. On modern paper machines the first open draw occurs in the press section, and therefore this is the point where most sheet breaks occur.

Modern paper machines often employ a cluster of three or four rolls that form two or three press nips, respectively. These multinip presses are more expensive to build and to operate, but they press the sheet in several nips before the first open draw. The repeatedly pressed sheet has a higher solids content and is stronger, which improves the paper machine runnability. Open draws have been eliminated from the wet end of some new paper machines of the future to permit higher operating speeds.

2.3 *Press Fabrics*

The role of press fabrics is to distribute evenly the pressure applied by grooved or drilled press rolls and to receive water expelled from the web in the press nip. In spite of its much heavier weight, the press fabric normally has a greater permeability than the wet web. Press fabric structures are carefully engineered to remain permeable to water even under compression. Modern press fabrics consist of a woven base covered by a batt of loose fibers. During its life, which might range from 10 days to 60 days, a press fabric passes through the press nip several million times. To prevent plugging of the fabric by various contaminants associated with the expelled water, continuous cleaning with showers, vacuum suction boxes, and, occasionally, chemical agents is required. When the fabric permeability can no longer be restored by chemical or other cleaning methods the fabric must be replaced.

2.4 *Mechanism of Pressing*

Water remaining in the fibrous web at the end of the forming section is located on the surface of the fibers, in the fiber lumens, and inside the swollen fiber walls. As the wet web and the adjacent press fabric start entering the press nip, the air contained in the sheet–fabric sandwich is expelled, and the nip becomes saturated with water. Further compression of the sheet and press fabric is possible only if some water escapes into cavities in the press roll. Water flow out of the sheet is driven by the hydraulic pressure in the press nip, but is limited by the low permeability of the sheet and its resistance to compression.

The compressed sheet and fabric, saturated with water in the midpoint of the press nip, begin to expand on the outgoing side of the nip. During this expansion, capillary forces transfer some of the water located within the press fabric back into the more compact wet paper in a process referred to as rewetting. Various authors have reported different estimates of the extent of rewetting in the outgoing side of the nip, but the real value is unknown. If the sheet is not separated from the press fabric immediately upon its exit from the nip, rewetting continues and can decrease the solids content of the pressed sheet by as much as 3%.

Numerous mathematical models with various degrees of complexity have been developed to describe water removal in the press section. Difficulties in modeling wet pressing include the short duration of the press pulse, the nonlinear variation in the permeability of the compressed sheet and press fabric, rewetting, and other factors. Models can be used to calculate the effect of pressing only by employing some experimentally derived factors, and the results are often not very accurate. Most models contain in some form the nip impulse, which is a product of nip pressure and the duration of pressure pulse. A practical form of the nip impulse is expressed by dividing the press nip load by the machine speed.

In addition to nip load and machine speed, the extent of pressing is strongly influenced by the sheet temperature. Increasing the temperature has a similar effect on the results of pressing as does increasing the press nip impulse, namely that the sheet solids content is higher and its thickness is reduced. On many machines, the sheet is heated directly at the entrance of the press nip by a steam shower or by infrared radiation. A heated press roll can also be used to heat the sheet, but this method is not commonly used. Considerable development of a technique commonly referred to as "impulse drying" has been carried out. In this approach the sheet is pressed using a press roll heated to a temperature around 300 °C. Impulse drying, which is more closely related to pressing than it is to drying, has yet to be used on a commercial machine. It is likely that the potential for improving pressing by further increasing the sheet temperature will be exploited in the future.

2.5 Problems Associated with Pressing

Pressing difficulties can be related to the press section operation or to the product quality. One common operational problem is press vibration, often caused by variations in the properties of roll covers or press fabrics. Sheet breaks can be caused by excessive sheet adhesion to the press roll at the outgoing side of the nip. Press fabric life can be drastically limited by incorrect cleaning or operating procedures. Unscheduled machine shutdowns, for premature replacement felts or press rolls, are a common cause of low machine efficiency.

While pressing of the wet web improves the quality of finished paper and board, it can also generate several types of defects. In particular, excessive pressing in the first nip can cause sheet crushing and holes in the product. Uneven moisture distribution across the paper width can be caused by unequal nip load distribution or by a variation in the press fabric quality. Sheet two-sidedness can arise when the side contacting the press fabric in the last press nip becomes rougher than the side contacting the smooth press roll.

3. Drying

In drying, the web is heated and moisture is removed by evaporation. The principles of heat and mass transfer are integral aspects of the process. The high energy requirements of the dryer section make it the most expensive part of the paper machine to operate, and paper is dried only when a large portion of water is removed by drainage, suction, and pressing. Typically, the web leaves the press section and enters the dryer section of a paper machine at a moisture content around 60%. The final moisture content after drying is 10% or less, depending on the grade and the desired properties (see *Paper: Effects of Moisture and Temperature**). During drying, bonding of the pressed and consolidated cellulosic fiber network occurs without the use of additional binders. Capillary forces developed during drying draw the fibers into close contact, providing in the fibrous network numerous sites for bonding (see *Paper: Structure**). The final product quality is greatly dependent on drying.

3.1 Multicylinder Dryers

The conventional multicylinder drying section used to dry most paper and paperboard is the longest part of the paper machine, sometimes having up to 90 cylinders. Paper is dried in a conduction-based process by contacting the hot surface of steam-heated cast iron cylinders with diameters of 1.2–1.8 m and shell thickness of 2.5–4 cm. The cylinders are pressure vessels typically rated for about 1.1 MPa. The pressure is gradually increased as the paper proceeds through the dryer. The steam is introduced into the cylinders through rotary joints, and condenses inside the cylinder. The condensate is evacuated from the cylinder through a rotary or stationary siphon along with some of the uncondensed steam, known as blow-through steam. The dryer is usually covered by a canopy or hood that may be partly or fully closed. This serves to exhaust the humid drying air and facilitates the recovery of energy for reuse. Hot, dry make-up air is continually supplied to the hood.

A typical cylinder dryer consumes 1.5 kg of steam per kilogram of water evaporated. The overall drying rate depends on the grade and the basis weight of the paper, and varies between 3 and 8 $\mathrm{g\,s^{-1}\,m^{-2}}$ for steam temperatures in the range 100–190 °C.

The cylinders are usually arranged in sections having upper and lower tiers, through which the paper travels in a serpentine fashion. The configuration, in which each tier has an endless felt or fabric to press the sheet against each dryer to maximize heat transfer, is known as a double-felted dryer section. The unsupported portion of the web is called a draw, and two neighboring draws together with the felt form a pocket. The sheet is not restrained in the draws, leaving it vulnerable to fluttering and breaks and allowing the sheet to shrink. The vapor released in the pocket induces localized high humidities that will retard further evaporation unless the pocket is ventilated with hot, dry air.

Over the years, incremental improvements to multicylinder drying have been primarily driven by the desire to increase the machine speed. Dryer felts were replaced by more permeable drying fabrics, and double-felted sections are being progressively replaced by single-felted ones, where a single drying fabric rather than an upper and lower fabric is used. Some single-felted dryer sections are built with a single tier of heated cylinders. Since the cylinders coming into contact with the fabric rather than the paper in a single-felted system provide little heating of the sheet, they are often replaced by plain or vacuum rolls. Single-tier sections are sometimes arranged in such a way as to provide alternate heating for both sides of the sheet.

The dryers for the fastest machines are now being designed as one long single tier with no alternating sections; thus one side of the sheet is always in contact with a fabric and the other in periodic contact with the drying cylinders. The sheet is heated from one side only. The single-tier arrangement improves machine runnability, which becomes critical as machine speeds approach 33 $\mathrm{m\,s^{-1}}$. Drying rate on some new machines is enhanced by blowing of hot combustion gas onto paper from a hood that surrounds the drying cylinder. This drying technique is expected to be widely used in the future.

It is important to achieve a uniform moisture profile across the width of paper. The web has a tendency to become wetter in the center, where the hood air is more humid than at the edges. If a uniform moisture profile is critical, the sheet can be overdried to a very low moisture level at the expense of operating efficiency. However, nonuniformities in moisture profile are usually corrected through the appropriate use of steam showers in the press section, remoisturizing showers in the dryer section, or infrared profiling units. Nonuniform moisture profiles cause difficulties in the subsequent operations such as coating, calendering, and winding, and may result in over dried or grainy edges, baggy or slack areas, and poor reel building. Other drying-related defects may include the following: blisters caused by an overheated cylinder; brittleness and loss of strength from over-drying; machine direction dryer wrinkles from a variety of causes, including misalignment, improper tension, or excessive flutter in the draw; felt seam marks or holes in the sheet from dripping of condensed vapor in the hood; and poor surface finish because of dirt on the cylinder.

3.2 Yankee Dryers

Drying of tissue and towel is carried out in contact with a single large polished cast iron cylinder called a Yankee dryer, which has a diameter of 3–6 m and a shell thickness of around 7.5 cm. Apart from their size and the special design considerations that this imposes, Yankee dryers are similar to the cylinders used in multicylinder drying and are heated with steam. Circumferential grooves inside the cylinder are used to provide more condensate heat transfer area, as well as decreasing the shell thickness through which heat must transfer before reaching the paper. Yankee dryers are almost always equipped with gas-fired impingement hoods or caps that greatly increase the drying rate by blowing jets of hot gas onto the paper surface. Jet velocities of around 200 $\mathrm{m\,s^{-1}}$ and jet temperatures up to 700 °C can be found on the most modern impingement hoods. Drying rates are around 560–1400 $\mathrm{g\,s^{-1}\,m^{-2}}$, which is one order of magnitude greater than those for multicylinder drying.

The wet web is pressed onto the hot dryer with a rubber roll, and the dry paper is detached from the dryer by means of a creping doctor. The microridges induced in the paper by the creping doctor create bulk and softness. Most Yankee dryers use a spray system to apply adhesion and release agents to the surface of the cylinder prior to the press roll. The proper adhesion and release of the sheet is critical to the drying and creping processes. Tissue and towel typically have a much lower basis weight than other paper grades, and are manufactured and dried at speeds beyond those of other paper products.

Machine glaze (MG) cylinders are similar in all respects to Yankee dryers except that the creping doctor is not used and the surface of the cylinder is more highly polished. The impingement hood is optional since the principal function of this type of dryer is to impart a smooth and glossy finish to one side of the paper or paperboard, rather than achieve a high drying rate. MG cylinders may be used in conjunction with conventional cylinder dryers, although more and more they are being replaced by special calendering processes that achieve the same finish and are simpler to operate.

3.3 Other Types of Dryer

Alternative drying methods using convective or radiation heat transfer, or a combination of pressing and drying, have found niche applications for certain paper grades. Coated papers requiring noncontact drying use combined convective and infrared drying systems. The quality of the coated surfaces is sensitive to the drying method, and the absorption of radiant energy in the infrared region is confined to a very thin layer especially suitable for coating drying. Sack grades requiring nonrestrained drying to induce product strength properties use air flotation drying. This technique uses high-velocity hot air streams from alternate air bars on both surfaces of the web for simultaneous floating and drying. Air flotation dryers are commonly used for drying thick sheets of market pulp. Microwave or dielectric drying is especially effective in correcting moisture profile for all grades.

Some tissue manufacturers use through-drying of an unpressed web to enhance softness and bulk. Hot air is drawn through the web, which is supported on a perforated cylinder under a vacuum. The drying rate is similar to that obtained on a Yankee dryer.

Much effort has been devoted to the development of press drying and impulse drying since the 1980s. In these concepts the distinction between pressing and drying is blurred since they are performed simultaneously to achieve high drying rates and special product properties. A few commercial designs have been applied in the manufacture of board grades.

There are now several applications using gas- rather than steam-heated cylinders in multicylinder dryers. This approach permits much higher surface temperatures than can be obtained with steam heating, and increases the drying rate.

3.4 Heat and Mass Transfer During Drying

Paper is dried at temperatures well below the boiling point of water. When wet paper is dried, simultaneous heat and mass transfer take place. For optimal drying, heat and mass transfer must be maximized while at the same time respecting the various operational and energy efficiency constraints. For multicylinder drying, the rate of heat transferred from the hot steam inside the cylinder to the cooler paper on the outside depends on the overall temperature gradient and the different resistances to heat transfer. These include the condensate inside the dryer, shell thickness, contact resistance at the paper–cylinder interface, and the web itself.

Increases in the drying rate can be achieved by increasing the heat transfer to the paper, which is the rate-limiting step for the greater part of the drying. This can be done by increasing steam pressure, which increases the overall temperature gradient, or by decreasing the resistances to heat transfer. The latter approach may include minimizing the thickness of the condensate layer rimming the inside of the cylinder through appropriate siphon arrangements, increasing the turbulence of the condensate layer through the use of longitudinal dryer bars inside the cylinder, keeping the inside and outside surfaces of the shell clean, and decreasing the contact resistance by increasing the fabric tension.

For Yankee dryers, heat transfer occurs not only from the steam-heated cylinder but also by convective heat transfer from the impinging jets. The contact resistance is low because the paper is pressed on the surface rather than being held with a fabric, and the heat transfer resistance of the web itself is low because of the low basis weights used for these grades.

After a sufficient amount of heat has been transferred to the wet web, a given mass of water vapor is transferred from the paper through a boundary layer to the air in the dryer section. This can occur by molecular diffusion, convective diffusion, and bulk movement or ventilation. The driving force for mass transfer is a vapor concentration (or partial pressure) difference between the web and the surrounding dryer air. The evaporation rate can be maximized by reducing the thickness of the boundary layer through proper pocket ventilation, keeping the vapor pressure of water in the sheet at a high level by keeping the sheet temperature elevated, and maintaining a low vapor pressure in the ambient dryer air by keeping it hot and dry.

3.5 Effects of Drying

For most applications, the rate of drying does not significantly affect the properties of paper, and the most important objective is to achieve a uniform moisture profile so that subsequent operations such as calendering, surface treatment, and printing will produce uniform results.

During drying, shrinkage of individual fibers in the radial direction induces shrinkage and cockle unless paper is restrained. Shrinkage begins to occur at about 65% solids content. The action of restraining the sheet during drying leads to the development of drying stress or tension, the extent of which depends on the degree of imposed restraints. Tension in the machine or cross-machine direction increases the tensile strength and reduces the toughness and stretch-to-failure in that direction. Restraint during drying reduces hygroexpansivity, and improves the dimensional stability of the sheet. Nonuniform drying stress in the thickness direction of the paper, as occurs in single-sided drying, causes the sheet to curl. Pressure in the thickness direction during drying increases the sheet density and lowers its porosity.

Multicylinder drying results in high machine direction tension although the sheet is relatively free to shrink intermittently in the cross-direction in the open draws. This cross-direction shrinkage is not uniform and is greater at the edges, resulting in nonuniform strength profiles. Air floatation drying used for sack grades allows shrinkage to occur in all directions with essentially no stress or strain on the sheet, resulting in high stretch and toughness but lower strength than is obtained in multicylinder drying.

4. Calendering, Winding, and Reeling

4.1 Calendering

Improving paper smoothness is the primary objective of calendering for printing and writing papers, and also for some packaging grades of paper and board. The dry web leaving the dryer section is bulky and has a rough surface. For most printing processes, print quality is proportional to paper smoothness. Even paper that is not used for printing or which is used as a base for coating is often calendered to improve roll building, since small differences in sheet thickness accumulate, causing variations in reel diameter and wrinkles in the sheet.

A calender consists of a vertical stack of 2–10 polished steel or chilled cast iron rolls. The paper passes in a serpentine fashion through the nips between each roll starting at the top of the stack. The pressure in each calender nip is generated by the weight of the rolls above the nip, or the calender might be hydraulically loaded or relieved. Linear nip loads do not usually exceed 150 kN m^{-1}. The paper is repeatedly compressed as it passes through each nip. After each compression, the paper expands but does not recover its original bulk. The surface of the calender roll is replicated onto the paper surface, and calender rolls must be periodically reground to maintain their smoothness as well as to correct any variations in their diameter caused by unequal wear.

Calendering is enhanced by all factors that render fibers less rigid and more malleable. Therefore, paper made from chemical pulp requires less calendering than paper made from mechanical pulps, and calendering can be improved by increasing the sheet moisture content and temperature. The effect of calendering is increased by increasing the number of nips or the nip load, and is decreased by increasing the machine speed or roll diameter.

The high nip loads in the calender can result in product damage. Calender blackening occurs at very high nip loads when fiber flocs are compressed to the extent that they lose their opacity. This defect occurs more readily if the sheet moisture content is high. Printing papers made from mechanical pulp may lose as much as 30% of their tensile strength during calendering, since compression of the sheet in the dry state causes destruction of fiber-to-fiber bonds. Creases in the paper entering the calender usually results in calender cuts. Any damage or scratch on the roll surface can be replicated on calendered paper.

In a supercalender, highly polished steel rolls alternate with rolls covered with a soft cover made from a plastic material or from heavily compressed cotton. The softness of this cover helps to prevent calender blackening, and therefore supercalenders can be loaded up to about 400 kNm^{-1}. Supercalenders are usually self-standing machines, used on coated paper and board to obtain high-gloss products.

Hot–soft calenders installed at the end of a paper machine can also be used to obtain gloss and smoothness values approaching those produced by supercalenders. These calenders usually consist of two or four nips formed by a metal roll heated to 160–200 °C and a "soft" roll with a synthetic cover.

4.2 Reeling and Winding

At the end of the paper machine, the dried and calendered sheet is reeled onto a large-diameter spool to produce large reels of paper called parent or jumbo rolls. The full machine width rolls obtained on the reel weigh several tonnes and are not suitable for final use. Therefore, a winder is used to unwind the jumbo rolls, slit the sheet, and convert the jumbo roll into smaller rolls suitable for shipment to press rooms or for further conversion to a final product. The basic principles of reeling and winding are similar.

During reeling, the spool of paper is pressed against a drum to obtain the desired roll density. The correct roll density profile is important for good final product quality. Roll density should be highest near its core and gradually decline toward the periphery of the roll. The density profile is controlled by the pressure between the spool and the drum and by the tension of the incoming paper. The top paper layers of the roll are under tension, and this belt of strained paper compresses the paper wound closer to the core. A well-wound roll is rigid and resists deformation by external forces.

Incorrect density profiles can result in various defects of the press roll, including bursts and excessive stretching of the outside plies of the roll. Rolls with too low a density near the core and a large paper tension at the perimeter can collapse in a star pattern when subjected to even modest external pressure. Therefore, reeling and winding are as important for the final product quality as any other papermaking operation.

5. Surface Treatment

The most common forms of surface treatment after drying involve application of sizing agents to increase the water repellency of the sheet and improve printability, barrier agents for packaging grades, and coating formulations required for various grades to improve surface optical and physical properties (see *Paper, Wetting and Penetration of Liquids into*[#]; *Pulp and Paper: Nonfibrous Components**; *Paper Products: Classification**).

Sizing agents such as starches and binders can be applied to dry paper or board using flooded nip, gate roll, or film transfer size presses. The degree of sheet rewet is minimized in film transfer size presses, requiring less sheet strength to prevent breaks and less after-drying. Precautions must be taken to prevent picking and sticking of the sheet on the first dryers following the size press.

Coating of the sheet with formulations containing pigments, minerals, binders, and other additives can be achieved off- or on-machine. Both sides of the sheet may be coated sequentially (with drying of the first side prior to coating of the second) or simultaneously after the calendering operation. Roll, blade, and metered film transfer coaters can be used to spread coating onto paper. Coatings are very delicate when wet, and initial drying must set the coating without disrupting it. For this reason, convective- and radiation-type dryers are used immediately after coating, although conventional dryers may be used after the initial drying.

See also: Paper Products: Classification*; Paper: Effects of Moisture and Temperature*; Paper: History of Development#; Paper: Pulping and Bleaching*; Paper: Strength and Stiffness*; Paper: Structure*; Paper, Wetting and Penetration of Liquids into#; Pulp and Paper: Nonfibrous Components*

Bibliography

Baker C F, Punton V W (eds.) 1989 *Fundamentals of Papermaking. Trans. 9th Fundamental Research Symp.* Mechanical Engineering Publications, London

Casey J P (ed.) 1980 *Pulp and Paper Chemistry and Chemical Technology*. Wiley, New York, Vol. 2

Gavelin N G (ed.) 1972 *Drying of Paper and Paperboard.* Lockwood, New York

Paulapuro H (ed.) 1999 *Papermaking Science and Technology, Vol. 8. Papermaking I. Papermaking II.* TAPPI Press, Atlanta, GA

Peel J D 1999 *Paper Science and Paper Manufacture*. Wilde, Vancouver, Canada

Polat O 1995 In: Mujumdar A S (ed.) *Handbook of Industrial Drying*. Dekker, New York

Reese D (ed.) 1999 *The Paper Machine Wet Press Manual*, 4th edn. TAPPI Press, Atlanta, GA

Robinson G W F, Baker C D (eds.) 1991 *Understanding and Troubleshooting the Paper Machine Dryer Section.* PAPTAC, Montreal, Canada

Smook G A 1987 *Handbook for Pulp and Paper Technologists.* Joint Textbook Committee of the Paper Industry, TAPPI, Atlanta, GA

Thorp B A, Kocurek M J (eds.) 1991 *Pulp and Paper Manufacture, Vol. 7. Paper Machine Operations.* Joint Textbook Committee of the Paper Industry, TAPPI, Atlanta, GA

N. A. Poirier, I. I. Pikulik and R. Gooding
Pulp and Paper Research Institute of Canada,
Pointe-Claire, Canada

PES and PEEK

The development of advanced polymer technologies is a crucial step in the application of biomedical polymeric materials, since common technologies require refinement in order to meet the stringent requirements for medical end use. Composite technology, plastics processing, and polymer synthesis are three well-known areas that, with adaptation, readily find use in biomedical applications. Biomaterials must survive the harsh chemical and mechanical environment of the body with minimal chance for revision once in place. Certain biomedical devices are expected to be biostable and have long-lasting properties in these harsh biological environments. Unlike in several other industries where a material is replaced after a defined time, the objective of the ideal implantable biomedical device is to be long lasting in a closed environment. Poly(ether ether ketone) (PEEK) and poly(ether sulfone) (PES), two advanced engineering thermoplastics, are examples of such materials and have been investigated for use in demanding medical applications.

1. Synthesis and Characteristics of PEEKs

PEEK is a linear, aromatic, thermoplastic, semicrystalline (about 20–30% crystallinity) material. It is hydrolytically stable, resistant to wear, stable in solvents, and abrasion resistant. Select properties are listed in Table 1. PEEK is synthesized by a condensation reaction (Fig. 1) and may be processed using standard melt-processing techniques such as roll forming, hydroforming, pultrusion, injection molding, and casting. Photodegradation of PEEK can occur with main chain scission and cross-linking occurring simultaneously below the glass transition temperature, while cross-linking largely occurs above the glass transition temperature. PEEK is commercially available in several grades according to different molecular weights (or melt viscosities) to meet the requirements of different melt-processing methods. It has been found that different grades of PEEK may be oriented in the solid state using compressive stresses to achieve uniaxial or biaxial orientation, thus yielding high modulus machinable stock for use in medical devices (Shalaby *et al.* 1996).

2. PEEK Applications: Past, Present, and Future

PEEK is commonly used in composites, both for biomedical and nonbiomedical applications, and it finds use as chromatography tubing or as wire and cable insulators. PEEK can be sterilized by autoclaving, by treating with ethylene oxide, or by high-energy radiation. Hollow fiber PEEK membranes can be produced by solution spinning. PEEK is a

Table 1
Representative properties of PES and PEEK.

Property	PES	PEEK
Density (g cm^{-3})	1.41	1.31
Tensile strength at yield (MPa)	99.4	90.0
Elongation to break (%)	30.2	39.2
Flexural yield strength (MPa)	140	150
Modulus of elasticity (GPa)	4	3.6
Impact strength, Izod (J cm^{-1})	0.8	0.74
Compressive yield strength (MPa)	100	120
Water absorption (%)	0.89	0.13
Deflection temperature at 1.8 MPa (°C)	210	170
Glass transition temperature (°C)	230	145
Melting temperature (°C)	350	330
Thermal conductivity (W m^{-1} K^{-1})	0.24	0.25
Electrical resistivity (Ω cm)	3.1×10^{16}	4.6×10^{15}
Dielectric constant	3.8	3.4

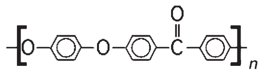


Figure 1
Molecular structure of PEEK.

promising candidate for hip stem manufacture, since its modulus is much lower than traditional metallic materials and it possesses much closer features to an isoelastic implant, as predicted by finite element models, thus minimizing potential for stress shielding-related problems most commonly associated with traditional metallic implants (Corvelli *et al.* 1996, 1997, 1999). PEEK is most commonly combined with glass or carbon fibers to form composite intramedullary devices. Select carbon fiber/PEEK systems, tested at various temperatures in solution up to 5000 hours, show no significant loss in compressive strength, modulus, and Poisson ratio as a function of time or temperature (Zhang *et al.* 1996). Carbon fiber/PEEK composite wear debris has been assessed in simulated physiological environments with a zirconia femoral head, PEEK wearing into particles of the order of 2.2 μm and carbon fiber into particles of the order of 153 μm. Carbon fiber/PEEK plates have been tested for *in vitro* biocompatibility and exhibit a nonspecific foreign body reaction only.

Biocompatibility tests of PEEK do not show signs of cytoxicity and they do show increased osteoblast cell protein content, which may be indicative of improved ingrowth performance (Morrison *et al.* 1995). Studies have also indicated that PEEK, processed in specific roughnesses, has relatively good biocompatibility, enhancing proliferation of osteoblastic cells and increasing production of osteocalcin (Lin *et al.* 1997). Attempts have been made to plasma spray hydroxyapatite onto PEEK/carbon fiber surfaces, a process commonly used with metallic implants to improve ingrowth and adhesion of an implant material. This method is not successful with thermoplastics, the hydroxyapatite in the PEEK composite systems demonstrating low shear strengths (HA) and fast debonding or delamination times. Hydroxyapatite may have application as a reinforcing agent for PEEK and preliminary studies have addressed the appropriate ratio and configurations of a composite using these two components (Abu Bakar *et al.* 1999).

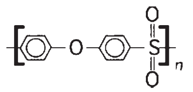


Figure 2
Molecular structure of PES.

Novel surface modification techniques have been developed in an effort to encourage osseointegration of PEEK implants with surrounding bone tissue. Two novel surface modification techniques have been employed to create microtextured surfaces, the first through microtexturing and the second by covalently bonding phosphonate groups (Shalaby and Rogers 1996, Shalaby and McCaig 1996, Shalaby and Roweton 1999). To achieve surface microtexturing with microporosity, the PEEK implants are subjected to so-called crystallization-induced microphase separation. To introduce phosphonate moieties on the PEEK surface that mimic a hydroxyapatite surface, PEEK is first acylated with propionic anhydride, then phosphonylated with PCl_3/O_2, and then hydrolyzed in the presence of aqueous calcium oxide. These investigations show that osseointegration can be achieved with transcortical tibial and mandibular implants in the rat and goat models, respectively.

3. *Aromatic PESs*

Polysulfones are engineering thermoplastics with desirable properties for use in biomedical applications. They are hydrolytically, thermally, and oxidatively stable and are unaffected by oil or grease. A commonly used PES is formed by a polycondensation reaction of a sulfonyl chloride and diphenyl ether to produce macromolecules with the repeat unit depicted in Fig. 2. Select properties of PES are listed in Table 1. PES is an amorphous thermoplastic material with high glass transition temperature. It can be obtained in a transparent grade and it can be rendered conductive. PES may be sterilized by autoclave, by treating with ethylene oxide, or by radiation. Compared to other amorphous polymers, it has one of the highest resistances to chemicals.

4. *PES Applications: Past, Present, and Future*

Typically, PES is used in the filtration industry as hollow fiber membranes as well as in respirator nebulizers. Due to its ability to survive high temperatures for sustained periods, PES is finding applications as an electrical insulator. Due to its grease and oil resistance, it also is used in the automotive and aircraft industries. PES has also found biomedical applications, where it is used as a resin material in composite intramedullary hip stems, material in instrument holders, reusable syringe injectors, prosthesis packaging, sterilizer trays, and dental tools. PES has been tested with varying boron ion doses to modulate the surface characteristics. Ion beam irradiation or glow discharge of its surface can improve cell adhesion (Pignataro *et al.* 1997, Tanahashi *et al.* 1996). Sliding tests of up to 10 000 cycles can actually increase the hardness of the polymer, in a manner proportional to the dose and beam energy (Rao *et al.* 1993). This therefore produces important wear characteristics, a critical issue for articulating implant surfaces. PES can be coated with apatite to improve its bioactivity (Nagano *et al.* 1996, Tanahashi *et al.* 1994, 1996) and bonding to bone. PES membranes have been used as hollow fibers in hemodialysis applications with some concern due to their thrombogenic features. Because of this property, the injection of an anticoagulent has become a necessity. New areas of interest therefore address blending the PES with other materials and thus improving its blood biocompatibility. PES has been tested as a hollow fiber encasing alginate and endothelial cell growth factor, with potential application to promote surrounding vasculature (Hunter *et al.* 1999). PES is used as dialysis membranes in order to attenuate reverse transfer of dialysate contaminants into the blood compartment.

5. *Comparisons of PEEK and PES*

Comparisons between select carbon fiber-reinforced PEEK and carbon fiber-reinforced PES show that the former maintain higher compressive and flexural strengths in simulated body environments (Strait *et al.* 1991). PEEK and PES with 30% short carbon fibers tested in saline show that select PEEK systems exhibit no change in flexion and fracture toughness while the PES system significantly decreases in mechanical properties with exposure (Brown *et al.* 1990). Interfacial bond strength in both microdroplet systems showed that the bond strengths significantly decrease following exposure to saline (up to 5000 hours) and are dependent on time and temperature. Following saturation, the immediate bond strength loss stabilizes at the temperature-dependent equilibrium bond strength level (Meyer *et al.* 1994). Biocompatibility studies of both matrix materials induce a low cellular response, indicating the promise of future use (Wenz *et al.* 1990). PES and PEEK grow competitive amounts of fibroblasts and osteoblasts as compared to traditional metallic materials and promote adhesion competitively (Hunter *et al.* 1995).

Bibliography

Abu Bakar M S, Cheang P, Khor K A 1999 Thermal processing of hydroxyapatite reinforced polyetheretherketone composites. *J. Mater. Processing Technol. Proc. 1999 4th Asia Pacific Conf. Materials Processing*, Elsevier, Lausanne, Switzerland, pp. 462–6

Brown S A, Hastings R S, Mason J J, Moet A 1990 Characterization of short-fibre reinforced thermoplastics for fracture fixation devices. *Biomaterials* **11** (8), 541–7

Corvelli A A, Bierma N N, Paul J. Roberts J C 1997 Finite element and experimental test results for PEEK composite intramedullary implants. In: *Proc. 1997 ASME Int. Mechanical Engineering Congress and Exposition*. ASME Bioengineering Division, Fairfield, NJ, pp. 313–14

Corvelli A A, Biermann P J, Roberts J C 1996 Design, analysis, and fabrication of a composite intermedullary implant. In: *Proc. 1996 11th Tech. Conf. American Society for Composites*. Technomic, Lancaster, PA, pp. 707–16

Corvelli A A, Roberts J C, Biermann P J, Cranmer J H 1999 Characterization of a peek composite segmental bone replacement implant. *J. Mater. Sci.* **34** (10), 2421–31

Hunter A, Archer C W, Walker P S, Blunn G W 1995 Attachment and proliferation of osteoblasts and fibroblasts on biomaterials for orthopaedic use. *Biomaterials* **16** (4), 287–95

Hunter S K, Kao J M, Wang Y, Benda J A, Rodgers V G 1999 Promotion of neovascularization around hollow fiber bioartificial organs using biologically active substances. *ASAIO J* **45** (1), 37–40

Lin T W, Corvelli A A, Frondoza C G, Roberts J C, Hungerford D S 1997 Glass PEEK composite promotes proliferation and osteocalcin production of human osteoblastic cells. *J. Biomed. Mater. Res.* **36** (2), 137–44

Meyer M R, Friedman R J, Del Schutte H Jr., Latour R A Jr 1994 Long-term durability of the interface in FRP composites after exposure to simulated physiologic saline environments. *J. Biomed. Mater. Res.* **28** (10), 1221–31

Morrison C, Macnair R, MacDonald C, Wykman A, Goldie I, Grant M H 1995 *In vitro* biocompatibility testing of polymers for orthopaedic implants using cultured fibroblasts and osteoblasts. *Biomaterials* **16** (13), 987–92

Nagano M, Kitsugi T, Nakamura T, Kokubo T, Tanahashi M 1996 Bone bonding ability of an apatite-coated polymer produced using a biomimetic method: a mechanical and histological study *in vivo*. *J. Biomed. Mater. Res.* **31** (4), 487–94

Pignataro B, Conte E, Scandurra A, Marletta G 1997 Improved cell adhesion to ion beam-irradiated polymer surfaces. *Biomaterials* **18** (22), 1461–70

Rao G R, Lee E H, Mansur L K 1993 Structure and dose effects on improved wear properties of ion-implanted polymers. *Wear* **162**, 739–47

Shalaby S W, Johnson R A, Deng M 1996 Process of making a bone healing device. *US Pat.* 5 529 736

Shalaby S W, McCaig S 1996 Process for phosphonylating the surface of an organic polymeric perform. *US Pat.* 5 491 198

Shalaby S W, Rogers K R 1996 Polymeric prosthesis having a phosphonylated surface. *US Pat.* 5558 517

Shalaby S W, Roweton S L 1999 Microporous polymeric foams and microtextured surfaces. *US Pat.* 5 969 020

Strait L H, Jamison R D, Gavens A 1991 Effect of environment on the flexural and compressive strength of carbon/polyetheretherketone composites. *Trans. Annu. Meet. Society of Biomaterials in Conjunction with the Int. Biomater. Symp.* Society of Biomaterials, Algonquin, IL, p. 286

Tanahashi M, Kokubo T, Nakamura T, Katsura Y, Nagano M 1996 Ultrastructural study of an apatite layer formed by a biomimetic process and its bonding to bone. *Biomaterials.* **17** (1), 47–51

Tanahashi M, Yao T, Kokubo T, Minoda M, Miyamoto T, Nakamura T, Yamamuro T J 1994 Apatite coated on organic polymers by biomimetic process: improvement in its adhesion to substrate by NaOH treatment. *J. Appl. Biomater.* **5** (4), 339–47

Wenz L M, Merritt K, Brown S A, Moet A, Steffee A D 1990 *In vitro* biocompatibility of polyetheretherketone and polysulfone composites. *J. Biomed. Mater. Res.* **24** (2), 207–15

Zhang G, Latour Jr. R A Kennedy J M, Del Schutte Jr. H, Friedman R J 1996 Long-term compressive property durability of carbon fibre-reinforced polyetheretherketone composite in physiological saline. *Biomaterials* **17** (8), 781–9

K. J. L. Burg[a] and S. W. Shalaby[b]

[a] *Clemson University, Clemson, South Carolina, USA*

[b] *PolyMed Inc., Anderson, South Carolina, USA*

Plastic Deformation of Cellular Materials

Foams made from materials which have a plastic yield point (rigid polymers, or metals, for instance) *collapse plastically* when loaded beyond the linear-elastic regime. Plastic collapse, like elastic buckling, gives a long horizontal plateau to the stress–strain curve, though the strain is not recoverable.

Plastic collapse in an open-cell foam occurs when the bending moment exerted on the cell walls exceeds the fully plastic moment (Thornton and Magee 1975, Gibson and Ashby 1997), creating plastic hinges like those shown in Fig. 1. Closed-cell foams are more complicated; in them, the plastic-collapse load may be affected by the stretching of cell faces as well as by the bending of the cell edges, and by the presence of a fluid within the cells. The simpler case of open cells is examined first.

1. Open-cell Foams

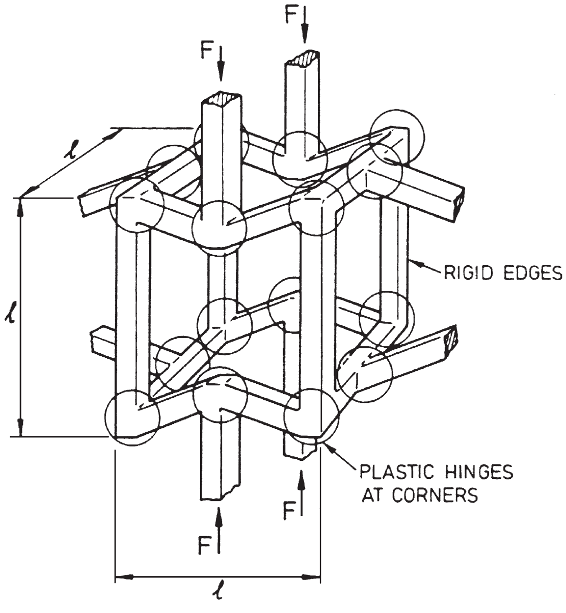


Figure 1
The formation of plastic hinges in an open-cell foam.

Plastic collapse occurs when the moment exerted by the force F (Fig. 1) exceeds the fully plastic moment of the cell edges. For a beam with a square section of side t, this fully plastic moment is:

$$M_{\mathrm{p}} = \tfrac{1}{4}\,\sigma_{\mathrm{ys}} t^3 \tag{1}$$

where σ_{ys} is the yield strength of the cell wall material. If the force F has a component normal to the cell edge (of length l), the maximum bending moment is proportional to Fl. The stress on the foam is proportional to F/l^2. Combining these results gives the plastic-collapse strength of the foam:

$$\sigma^*_{\mathrm{pl}} \propto \frac{M_{\mathrm{p}}}{l^3}$$

For open-cell foams, the relative density $\rho^*/\rho_{\mathrm{s}} \propto (t/l)^2$ (where ρ^* is the density of the foam and ρ_{s} is that of the solid of which it is made) giving the collapse stress of the foam σ^*_{pl}:

$$\frac{\sigma^*_{\mathrm{pl}}}{\sigma_{\mathrm{ys}}} = C_1\left(\frac{\rho^*}{\rho_{\mathrm{s}}}\right)^{3/2} \quad \text{(open cells)} \tag{2}$$

where the constant C_1 contains all the constants of proportionality. At higher densities the dimensions of the cell corners must be subtracted from the length of the beam, and corrections must be made to the equation for the density. When this is done (Gibson and Ashby 1997) the last equation becomes:

$$\frac{\sigma^*_{\mathrm{pl}}}{\sigma_{\mathrm{ys}}} = C_1'\left(\frac{\rho^*}{\rho_s}\right)^{3/2}\left(1+\left(\frac{\rho^*}{\rho_{\mathrm{s}}}\right)^{1/2}\right) \tag{3}$$

(open cells, including density correction)

Two limits on this behavior are worth noting. At large relative densities ($\rho^*/\rho_s > 0.3$), the beam-bending concept breaks down. The cell walls are now so short and squat that they yield axially (in compression or tension) before they bend. Above this density the material is better thought of as a solid with holes in it, not as a foam. At the other extreme—that of very low relative density—it is possible for elastic collapse to precede plastic collapse, even though the subsequent large-strain behavior (and thus the plateau stress) may be determined by plasticity.

2. *Comparison with Data for Open-cell Foams*

Data for the plastic-collapse strength and for the plastic indentation of foams (discussed further on) are plotted in Fig. 2. Open symbols are for open cells; full symbols are for closed cells. They are adequately described by Eqn. (2) with $C_1 = 0.3$ (full line, Fig. 2); Eqn. (3) with $C_1' = 0.23$ does marginally better (broken line). Each set of data follows the theory well, but with differing values for C_1, probably because of the errors in the value taken for yield strength of the solid, which is used as a normalizing parameter. In summary:

$$\frac{\sigma^*_{\text{pl}}}{\sigma_{\text{ys}}} \approx 0.3\left(\frac{\rho^*}{\rho_s}\right)^3 \quad \text{(open cells)} \tag{4a}$$

or

$$\frac{\sigma^*_{\text{pl}}}{\sigma_{\text{ys}}} \approx 0.23\left(\frac{\rho^*}{\rho_s}\right)^{3/2}\left(1+\left(\frac{\rho^*}{\rho_s}\right)^{1/2}\right) \tag{4b}$$

(open cells, including density correction)

3. *Closed-cell Foams*

Closed cells have membranes spanning their faces. Plastic collapse causes the membranes to fold and crumple in the compression direction; being thin, the force required to crumple them is small. But at right angles to this direction the membranes are stretched, and the plastic work required to extend them contributes significantly to the yield strength of the foam. This is assessed by calculating the work done in an increment of deformation and equating it to the plastic dissipation in bending the cell edges and stretching the cell faces.

Consider the collapse of a cell like that shown in Fig. 3. The cell edges and faces have thicknesses t_e and t_f, respectively; they are related to the fraction of solid in the edges, ϕ, by:

$$\frac{t_f}{l} = 1.41(1-\phi)\frac{\rho^*}{\rho_s} \tag{5a}$$

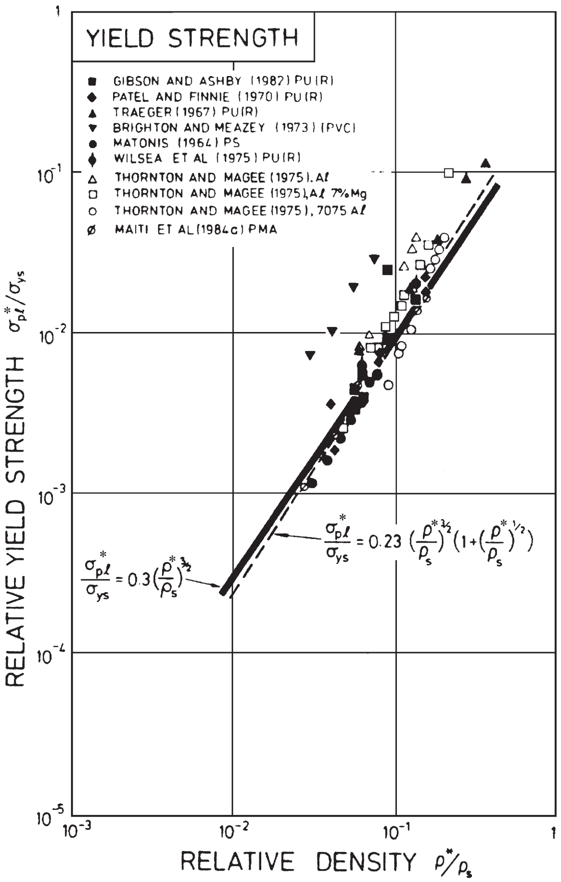


Figure 2
Data for the compressive yield strength of foams, normalized by the yield strength of the solid, plotted against relative density. The solid line represents the theory for open-cell foams; the dashed line represents that for open-cell foams with an additional density correction. They are almost indistinguishable.

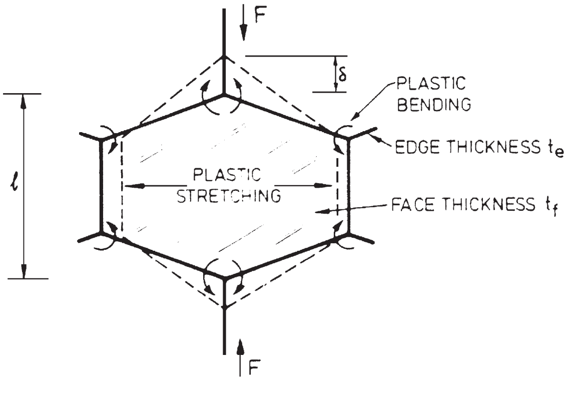


Figure 3
Plastic stretching of the cell faces of a closed-cell foam.

and

$$\frac{t_{\mathrm{e}}}{l} = 0.93\phi^{1/2}\left(\frac{\rho^*}{\rho_{\mathrm{s}}}\right)^{1/2} \tag{5b}$$

(Gibson and Ashby 1997).

A compressive plastic displacement δ of one cell allows the applied force F to do work $F\delta$. The angle of rotation at the four plastic hinges is then proportional to δ/l, and the plastic work done at these hinges is proportional to $M_{\mathrm{p}}\delta/l$. The cell face is stretched by a distance which is again proportional to δ, doing work which scales as $\sigma_{\mathrm{ys}}\delta t_{\mathrm{f}}l$. Summing these and equating to $F\delta$ gives:

$$F\delta = \alpha M_{\mathrm{p}}\frac{\delta}{l} + \beta\sigma_{\mathrm{ys}}\delta t_{\mathrm{f}}l$$

where α and β are constants. Replacing F by σl^2, and M_{p} by $\sigma_{\mathrm{ys}}t_{\mathrm{e}}^3/4$ (Eqn. (1) with $t = t_{\mathrm{e}}$) gives:

$$\frac{\sigma^*_{\mathrm{pl}}}{\sigma_{\mathrm{ys}}} = \frac{\alpha}{4}\left(\frac{t_{\mathrm{e}}}{l}\right)^3 + \beta\left(\frac{t_{\mathrm{f}}}{l}\right).$$

Finally, using Eqns. (5a) and (5b) we obtain:

$$\frac{\sigma^*_{\mathrm{pl}}}{\sigma_{\mathrm{ys}}} = C_1\left(\phi\frac{\rho^*}{\rho_{\mathrm{s}}}\right)^{3/2} + C_1''(1-\phi)\left(\frac{\rho^*}{\rho_{\mathrm{s}}}\right) \tag{6}$$

(closed cells, including membrane stresses) where C_1 and C_1'' contain the earlier constants. In the limit of $\phi = 1$ (open cells), we already know that $C_1 = 0.3$; in the other limit of $\phi = 1$, σ^*_{pl} must become equal to σ_{ys} when $\rho^* = \rho_{\mathrm{s}}$, so that $C_1'' \approx 1$. A density correction factor can be applied to the first term if desired (see "open cells," above), but its influence is small.

Equation (6) describes the combined effect of the plastic bending of cell edges and plastic stretching of their faces. To it must be added the contribution of the fluid contained in the cells. If the cell fluid is a gas at atmospheric pressure, p_{at}, the contribution is usually small (because $\sigma_{\mathrm{ys}} \gg p_{\mathrm{at}}$). But if the cells contain fluid at a pressure p_0 much larger than atmospheric, this pressure must be added to the right-hand side of Eqn. (6).

4. Comparison with Data for Closed Cells

The limited data for the plastic collapse of closed-cell foams are plotted in Fig. 2 as solid symbols. The scatter is considerable, due largely to uncertain values of σ_{ys}. The data are adequately described by Eqn. (4a); there is no evidence here for the influence of membrane stresses or of cell fluid. This may be because the cell faces rupture before full plastic collapse: then the closed-cell foam will behave like one with open faces. Or it may be because, in these foams, $\phi \approx 1$, and the contribution of the faces, though present, is small. It should not, in our judgement, be ignored even though direct evidence for it is lacking. The best current description of the collapse strength of closed-cell, plastic foams is then:

$$\frac{\sigma^*_{\mathrm{pl}}}{\sigma_{\mathrm{ys}}} \approx 0.3\left(\phi\frac{\rho^*}{\rho_{\mathrm{s}}}\right)^{3/2} + (1-\phi)\frac{\rho^*}{\rho_{\mathrm{s}}} + \frac{p_0 - p_{\mathrm{at}}}{\sigma_{\mathrm{ys}}} \tag{7}$$

(closed cells, including membrane stresses and gas pressure contribution).

5. Densification

Large plastic strains in compression cause the cell walls to crush together, and make the stress–strain curves rise steeply to a limiting strain ε_D. To an acceptable approximation,

$$\varepsilon_D = 1 - 1.4\left(\frac{\rho^*}{\rho_{\mathrm{s}}}\right) \tag{8}$$

6. Plastic Indentation

Unlike dense solids, which do not change their volume when deformed plastically, foams compact when compressed. The cells of the foam collapse as the foam is squeezed, so that axial compression produces very little lateral spreading once collapse has begun: the ratio of the lateral spreading to the axial compression is, typically, 0.04.

When a foam is compressed by a flat punch, the cells beneath the punch collapse in the direction of punching, but expand sideways hardly at all. Because of this the material in the plastic zone is not constrained by its surroundings in the way that the zone in a dense solid would be. An analysis of this problem

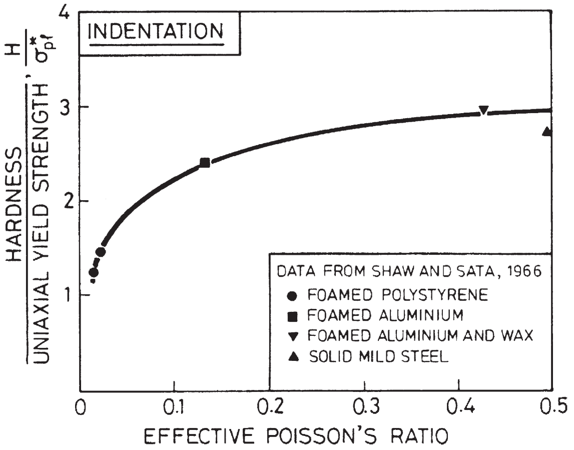


Figure 4
The ratio of indentation hardness to the uni-axial yield strength as a function of the effective plastic Poisson's ratio during plastic yielding.

(Wilsea *et al.* 1975) shows that, for relative densities less than about 0.3, the indentation pressure, or "hardness," H, of the foam is simply:

$$H = \sigma^*_{\mathrm{pl}} \qquad (9)$$

(instead of the value $H = 3\sigma_{\mathrm{ys}}$ which characterizes a dense solid). The result is obvious: the lack of sideways expansion makes the indentation test exactly like a compression test. As the effective Poisson's ratio (the negative ratio of the lateral strain to the axial compressive strain) increases, the hardness increases from σ^*_{pl} to $3\sigma^*_{\mathrm{pl}}$, as shown in Fig. 4.

See also: Foams, Microrheology of*; Plastic Deformation: Constitutive Description#

Bibliography

Ashby M F, Evans A G, Fleck N A, Gibson L J, Hutchinson J W, Wadley H N G 2000 *Metal Foams: A Design Guide*. Butterworth Heinemann, Oxford, UK

Brighton C A, Meazey A E 1973 *Expanded Plastics*. QMC Industrial Research Unit, London

Gibson L J, Ashby M F 1997 *Cellular Solids: Structure and Properties*, 2nd edn. Cambridge University Press, Cambridge

Maiti S K, Gibson L J, Ashby M F 1984 *Acta. Metal.* **32**, 1963

Matonis V A 1964 *J. Soc. Plast. Eng*. September issue, 1024

Patel M R, Finnie I 1970 *J. Mater.* **5**, 909

Shaw M C, Sata T 1996 *Int. J. Mech. Sci.* **8**, 469

Thornton P H, Magee C L 1975 *Met. Trans.* **6A**, 1253

Traeger R K 1967 *J. Cell. Plast.* **3**, 405

Wilsea M, Johnson K L, Ashby M F 1975 *Int. J. Mech. Sci.* **17**, 457

M. F. Ashby
University of Cambridge, UK

Polymer Composites, Friction and Wear of

Polymer composite materials, one of the most rapidly growing classes of materials, are being used increasingly for technical applications, especially where their low density and high specific strength and stiffness are of great interest. Another advantage is that their properties can be varied and optimized over a wide range with regard to the requirements of the respective technical problem. Polymeric materials containing different fillers and/or reinforcements are also being used increasingly in different tribological applications (Friedrich 1986, 1993, Lancaster 1983, Zum Gahr 1987).

Some of these applications, such as sliding elements, require low friction and wear, whereas for other applications, for example in clutches or brakes, high friction combined with low wear is necessary. In all cases friction coefficient and wear rate are not individual material properties but are properties of the whole tribological system in which the materials function. This means that the tribological behavior of a material is operatively determined by the microstructure and the surface morphology of the materials in contact, the stress and speed conditions involved, and many other parameters of the tribosystem employed (Czichos and Habig 1992, Uetz and Wiedemeyer 1985). In such applications, composite materials are of particular interest since these materials contain highly tailored combinations of materials, the nature, geometry, and spatial distributions of which can be designed with considerable freedom, and controlled with considerable precision in processing.

1. Tribological Properties of Polymer Composites

The polymer-based composite–steel tribological system is of particular interest for technical use. In this context there are relatively few polymers that have valuable tribological properties, and most research is focused on these. Common thermoplastic polymers with actual or potential tribological function are listed in Table 1 together with their basic mechanical characteristics (Stachowiak and Batchelor 1993).

In general, the friction and wear behavior of polymeric matrix materials can be improved by a lower surface tension and a higher stiffness and strength. This can be achieved quite successfully by using special fillers. To reduce the adhesion, internal lubricants, such as PTFE or graphite, are frequently incorporated. Short aramid, carbon, or glass fibers are used to increase stiffness, strength, impact resistance, thermal conductivity, and creep resistance of the polymer system. In this way, excellent friction and wear properties can be achieved by a composite material consisting of a high-performance polymer matrix with lubricants and fibrous reinforcement (Häger and Davies 1993, Reinicke *et al.* 1998).

2. Wear Mechanisms

For a comprehensive understanding of the tribological behavior of polymer composite materials a close look at the possible wear mechanisms is important. Quite often the latter are subdivided into adhesive, abrasive, surface fatigue, and tribochemical mechanisms (Czichos and Habig 1992). For fiber-reinforced polymer composites a further classification of the various wear features is useful (Fig. 1). Besides material removal due to fiber and matrix sliding wear, a great amount of wear debris is generated by fiber cracking and fiber/matrix interfacial separation (Friedrich 1993). For unidirectional fiber-reinforced polymer composite under normal fiber orientation

Table 1
Tribological characteristics of selected thermoplastic polymers (Stachowiak and Batchelor 1993).

Polymer	Tribological characteristics	Temperature limit (°C)	Tensile modulus (GPa)	Tensile strength (Mpa)
Polytetrafluorethylene (PTFE)	Low friction, high wear Reinforcing component necessary High operating temperature limit	260	0.7	25
Polyamides (PA11/12, PA6/66)	Moderate to high friction, low wear Wear accelerated by water Moderate temperature limit	80–130	1.8–3.3	50–90
Polyacetals (POM)	Performance similar to polyamide Durable in rolling contacts High friction, low wear	100	3.3	70
Polyetheretherketone (PEEK)	Very high load bearing capacity High operating temperature limit Moderate friction, low wear	250	3.6	97
Polyethylenes (e.g. UHMWPE)	Low wear also in water Low operating temperature limit Moderate friction, low wear	100	0.6	20–40
Polyimides (PI, PAI, PEI, PBI)	Very high load bearing capacity High operating temperature limit	170–300	2.8–4.9	95–190

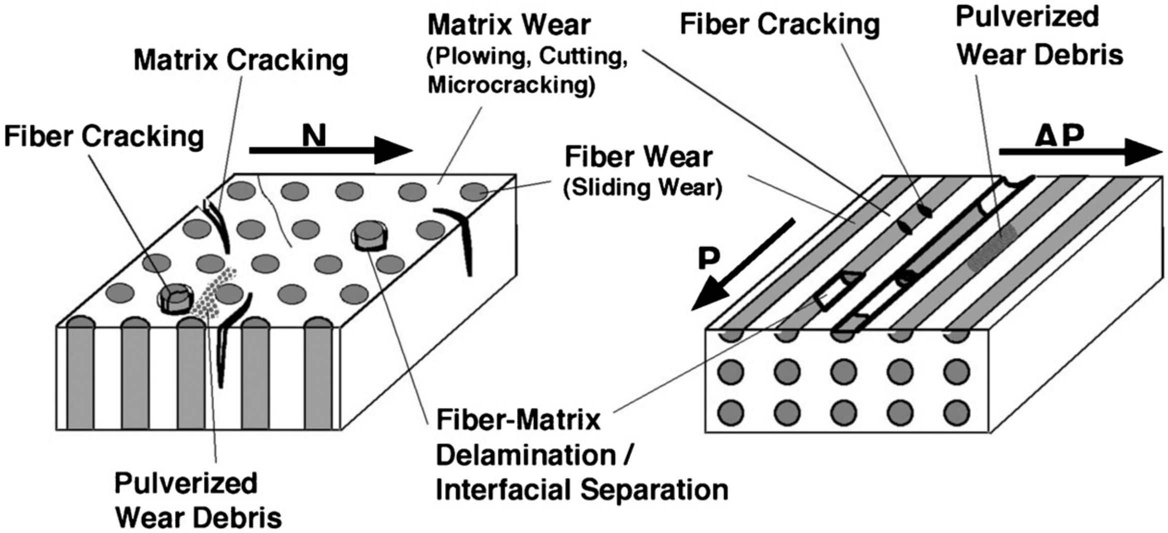

Matrix Cracking
Matrix Wear
(Plowing, Cutting, Microcracking)
Fiber Cracking
Pulverized Wear Debris
Fiber Cracking
N
ΔP
Fiber Wear
(Sliding Wear)
P
Fiber-Matrix Delamination / Interfacial Separation
Pulverized Wear Debris

Figure 1
Wear mechanisms in composite materials (Friedrich 1993).

(N), where fiber ends are protruding at the sliding surface, a possible sequence of wear steps is that the softer polymer matrix between the much harder fibers is removed first, followed by cracking of fiber pieces, leading to a repetition of the wear cycle. It should also be mentioned here that the whole process can be superimposed by the formation of cushioning layers of consolidated wear debris, by which the surface contact conditions are modified and the resulting wear processes frequently alleviated (Williams *et al.* 1997).

According to variations in wear mechanisms different fiber orientations relative to the sliding motion can cause a drastic change in the wear behavior of composite systems. In principle, three ideal fiber orientations exist. Beside the normal (N) fiber orientation mentioned already, parallel (P) and antiparallel (AP) fiber orientations, with fibers laying in the contact plane, exist (Fig. 1). The effect of fiber orientation on the tribological behavior also depends strongly on the type of fiber employed. Experimental studies on the interaction of various fiber-reinforced polymer composites with steel counterparts have shown that, depending on the type of fiber (aramid (AF), carbon (CF), glass (GF)), the optimum fiber orientation relative to the sliding direction differs. Aramid fibers, which also interact less abrasively against the steel surface (Friedrich 1993), exhibit, due to their tendency for fibrillation and their poorer bonding to the polymer matrix, a better performance under N-orientation, whereas carbon fibers have proved to be most beneficial for wear resistance of the composite under P-orientation. The reason for the higher wear rate for CF-composites under N- relative to P- or AP-orientations is based on the fact that the upper load limits (strength) of the fibers (under the same external load enhancement) are reached inside the composite earlier under N-orientation (Friedrich *et al.* 1999, Varadi *et al.* 1999). Another effect restricting the applicability of CF-composites with normal fiber orientation is a severe wear or seizure of the counterpart that can be initiated by exposed fiber ends.

3. *Tribological Behavior of Short-fiber Reinforced Polymers*

To characterize the tribological behavior of polymer composite materials standard tests are used based on the tribotechnical system under consideration. The pin-on-disk test, one of the most frequently used test configurations, is described in detail in Czichos and Habig (1992) and in Uetz and Wiedemeyer (1985). The results of such a pin-on-disk test on three high-temperature-resistant polymers (polyethernitrile (PEN), polyetheretherketone (PEEK), polyetheretherketoneketone (PEEKK)) with different internal lubricants and short fibers are shown in Fig. 2 (Friedrich *et al.* 1995, Lu and Friedrich 1997).

This figure illustrates for two different testing temperatures that the fibers primarily reduce the wear rate, whereas the internal lubricants, such as polytetrafluoroethylene (TF) and graphite (Gr), result in a reduction of the coefficient of friction. According to both the coefficient of friction and the specific wear rate, a system consisting of a PEEK-matrix with 10% each of CF, TF, and Gr seems to be the best compromise. In the temperature range under consideration (20–50 °C) a low coefficient of friction ($\mu < 0.2$)

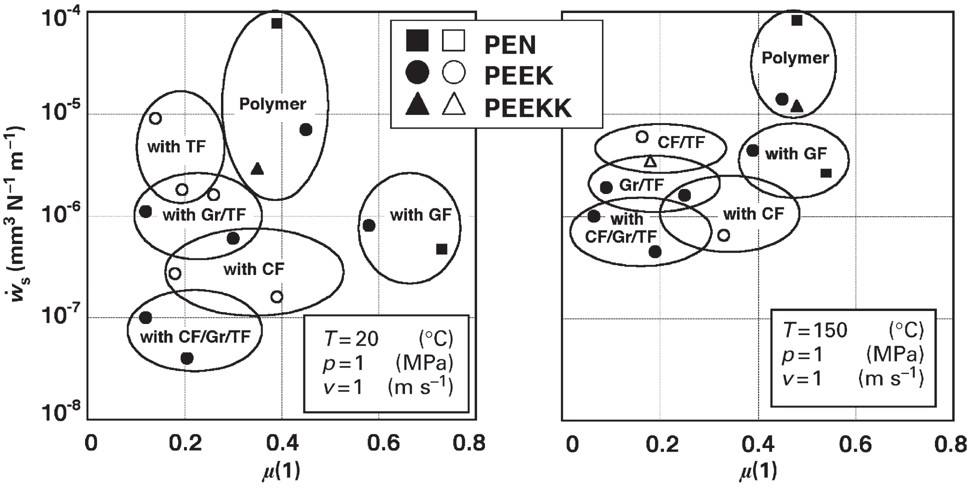


Figure 2
Tribological behavior of PEN, PEEK, and PEEKK composites (same symbols in one circle refer to samples with different filler volume fraction) (Lu and Friedrich 1997).

and a low wear rate ($\dot{w}_s < 1.5 \times 10^{-6}\,\mathrm{mm^3 (N\,m)^{-1}}$) can be achieved for an applied pressure $p = 1\,\mathrm{MPa}$ and a sliding velocity $v = 1\,\mathrm{m\,s^{-1}}$.

When considering the choice of a certain composite material for technical applications, two questions are of concern: (i) Which material can bear the load considered? (ii) How can a given component be produced from this material? Although the first question is of primary concern, the question of economics will ultimately lead to the choice of a particular material (Hornbogen 1993). For technical applications there are always several requirements, a so-called requirement profile, to be fulfilled. Within the scope of the requirement profile, there are properties of primary concern e.g., a low coefficient of friction and a low wear rate for sliding elements. Secondly, high stiffness and strength could be required. Furthermore, the demand for high hardness is often combined with the requirement for sufficient fracture toughness. As an example, consider sliding shoes attached to heavy metallic chains in textile drying machines. These must operate under extreme conditions, including elevated temperature and impact events, so an injection-moldable polymer composite with high fracture toughness as well as low friction and wear is required. An evaluation of various potential material systems (PI, PBI, PEEK) leads to the final choice of a short-fiber (fiber length $\sim 300\,\mu\mathrm{m}$) reinforced PEEK matrix composite (including 10% each of carbon fibers, PTFE particles, and graphite flakes) (Lu and Friedrich 1997). In particular, the PEEK system performs quite well within a wide range of the pressure (p) × velocity (v) product, at various testing temperatures. It also exhibits a high fracture toughness and is injection moldable into complex geometries.

4. *Tribological Behavior of Continuous-fiber Reinforced Polymers*

For applications requiring extremely high stiffness and strength, high-performance polymer composites are frequently used. The main characteristics of these composites are a continuous (endless) fiber reinforcement, a high fiber content (up to 60 vol.%), and a single (unidirectional) fiber orientation. For a systematic development of such a high-performance composite with regard to an optimum wear resistance, various unidirectional composites have been tested with respect to their wear behavior as a function of matrix material, type of fiber reinforcement, and orientation of fibers relative to the sliding direction. Several matrices, comprising one rather brittle one (epoxy resin) and two very tough thermoplastics (PEEK, PA66), and three different fiber materials (glass, carbon, and aramid) were chosen for this investigation (Table 2, Friedrich 1993, Cirino *et al.* 1988). The wear resistance ratios are normalized to that of neat EP ($\dot{w}_s^{-1} = 1.8 \times 10^4\,\mathrm{N\,m\,mm^{-3}}$), so a higher ratio means a better wear resistance relative to the EP standard.

The results of Table 2 indicate that:

(i) Tougher matrices can quite often be expected to show a higher wear resistance than brittle ones.

(ii) In a matrix with a lower wear rate, the beneficial effect of the fiber reinforcement on the wear resistance of the composite is not as pronounced as in an easily wearing matrix.

(iii) For most of the fiber orientations, carbon fibers give better results than glass fibers, while aramid fibers are somewhere in between.

(iv) As mentioned before, in the case of carbon fiber reinforcement, higher values of the wear rate are measured for N-orientation than for P- and AP-orientations. The opposite is true for aramid fibers.

The wear resistance (inverse of the wear rate) of these materials can be utilized to design a composite material with a generally good wear performance. Optimum wear resistance can be expected from a composite with a PEEK matrix and carbon fibers parallel to the sliding direction. If multidirectional sliding is the case in practical application, additional carbon fibers in antiparallel orientation will probably be helpful. In fact, a study of such a material, PEEK

Table 2
Wear resistance of UD composites (normalized to neat EP, $\dot{w}_s^{-1} = 1.8 \times 10^4\,\mathrm{N\,m\,mm^{-3}}$).

Matrix	Orientation	CF	GF	AF	Neat matrix
EP	N	65.4	0.5	122.4	1
	P	99.8	2.1	44.1	1
	AP	44.1	1.1	46.3	1
PEEK	N	65.1	5.1	61.4	11.8
	P	158.1	5.5	46.1	11.8
	AP	94.8	2.0	44.1	11.8
PA66	N	57.0	1.8	145.5	27.7
	P	122.9	16.3	60.8	27.7
	AP	80.2	8.9	73.7	27.7

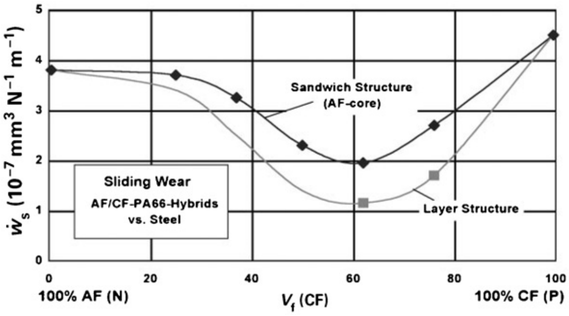


Figure 3
Synergistic effects for wear resistance of 2D hybrid composites with normal (aramid fibers) and parallel (carbon fibers) orientation.

with carbon fiber woven fabric reinforcement, resulted in even better wear factors than those of the two basic orientations of the unidirectional material. This synergistic effect is probably a result of fiber interlocking in the contact area due to the woven structure of the reinforcement.

A composite with good overall wear resistance can be made by three-dimensional (3D) hybridization, with interwoven carbon fibers in the *xy*-plane and aramid fibers in the *z*-direction. The matrix should be a very tough, highly wear-resistant polymer such as PEEK. Although such a 3D hybrid structure could not be verified by the studies of Cirino *et al.* 1988, other works with simpler 2D hybrid composites point in the direction predicted. Krey *et al.* 1989 have manufactured different laminates from amorphous polyamide-impregnated carbon and aramid strands (J2-Polymer, Du Pont, Wilmington, USA). Sliding these materials against rotating steel rings resulted in synergistic effects of the wear rates of hybrid composites with aramid fibers under normal and carbon fibers under parallel orientation (Fig. 3).

There are also some indications in the literature that composites with 3D fiber arrangement and a high-temperature polymer matrix with lubricants as inclusions have excellent wear resistance. In the particular case mentioned, glass or carbon fibers were used as reinforcements. The 3D structure was studied for different geometries of the sliding elements e.g., as flat blocks and circular bearings produced by 3D weaving technology with fibers in radial, tangential, and axial directions (Gardos 1986).

5. *Conclusions*

The study and examination of the tribological behavior of fiber-reinforced polymer composite systems has become an important contribution to the general field of tribology. Besides the general choice of the materials used for the composite system, a crucial parameter affecting the wear performance is the fiber orientation relative to the sliding direction. Therefore, a basic understanding of the different wear mechanisms involved in the wear process is essential for the interpretation of the tribological behavior of composites.

Interest in polymers and polymer-based composites for technical applications has grown rapidly. Fiber-reinforced high-temperature-resistant thermoplastic materials can now be used as sliding elements where only metallic components were used before. For high-performance thermoplastic matrix systems applications are limited to temperatures of up to about 250 °C depending on the loading conditions employed. Improvements in the tribological properties of these composites are usually attained through the incorporation of different fillers as internal lubricants. By the specific choice of material composition, it is possible to tailor the performance of polymer composites according to the special demands of different requirement profiles. Further improvements in the tribological behavior can be achieved by optimizing the design of continuous-fiber reinforced, high-performance composites. A three-dimensional arrangement of aramid and carbon fibers with different fiber orientations relative to the sliding direction gives a strong improvement in the wear behavior through synergistic effects between the two fibers. For these composites the load limits, which ultimately determine the limits of application, can be further increased. In conclusion it can be stated that advances in the tribological behavior of polymer composites offer new arrays of potential applications. One innovative example is the development of functionally graded tribomaterials, in which the wear resistance along with other properties varies over the cross-section of a part due to intentional variations in the microstructure of the material (Klingshirn *et al.* 2000).

Bibliography

Cirino M, Friedrich K, Pipes R B 1988 Evaluation of polymer composites for sliding and abrasive wear applications. *Composites* **19**, 383–92

Czichos H, Habig K-H 1992 *Tribologie Handbuch*. Vieweg, Braunschweig, Germany

Friedrich K 1986 Preface. In: Friedrich K (ed.) *Friction and Wear of Polymer Composites*. Elsevier, Amsterdam, pp. v–vii

Friedrich K 1993 Wear models for multiphase materials and synergistic effects in polymeric hybrid composites. In: Friedrich K (eds.) *Advances in Composite Tribology*. Elsevier, Amsterdam, pp. 209–73

Friedrich K, Flöck J, Varadi K, Neder Z 1999 Numerical and finite element contact and thermal analysis of real composite-steel surfaces in sliding contact. *Wear* **225–9**, 368–79

Friedrich K, Lu Z, Häger A M 1995 Recent advances in polymer composites: tribology. *Wear* **190**, 139–44

Gardos M N 1986 Self-lubricating composites for extreme environmental conditions. In: Friedrich K (eds.) *Friction* *and Wear of Polymer Composites*. Elsevier, Amsterdam, pp. 397–447

Häger A M, Davies M 1993 Short-fiber reinforced, high-temperature resistant polymers for a wide field of tribological applications. In: Friedrich K (eds.) *Advances in Composite Tribology*. Elsevier, Amsterdam, pp. 107–57

Hornbogen E 1993 *Werkstoffe*. Springer, Berlin

Klingshirn C, Koizumi M, Haupert F, Giertzsch H, Friedrich K 2000 Structure and wear of centrifuged epoxy-resin/carbon fiber functionally graded materials. *J. Mater. Sci. Lett.* **19**, 263–6

Krey J, Friedrich K, Jakobs O 1989 Herstellung und Untersuchung der Verschleißeigenschaften von Hochleistungsverbundwerkstoffen mit Thermoplast-Matrix. *VDI-Ber.* **737**, 179–94

Lancaster J K 1983 Composites for increased wear resistance: current achievements and future prospects. In: *NASA Conference 2300: "Tribology in the 80's."* NASA, Cleveland, OH, pp. 333–56

Lu Z, Friedrich K 1997 Polymere Hochtemperatur-Verbundwerkstoffe für Anwendungen als Gleitelemente. *Materialwiss. Werkstofftech* **28**, 116–23

Reinicke R, Haupert F, Friedrich K 1998 On the tribological behaviour of selected, injection moulded thermoplastic composites. *Composites* **29A**, 763–71

Stachowiak G W, Batchelor A W 1993 *Engineering Tribology*. Elsevier, Amsterdam

Uetz H, Wiedemeyer J 1985 *Tribologie der Polymere*. Hanser, München, Germany

Varadi K, Neder Z, Friedrich K, Flöck J 1999 Finite element analysis of a polymer composite subjected to ball indentation. *Compos. Sci. Technol.* **59**, 271–81

Williams J A, Morris J H, Ball A 1997 The effect of transfer layers on the surface contact and wear of carbon-graphite materials. *Tribol. Int.* **30**, 663–76

Zum Gahr K H 1987 *Microstructure and Wear of Materials* Elsevier, Amsterdam

K. Friedrich and R. Reinicke
Universität Kaiserslautern, Germany

Polymer Fiber Processing: High-performance Fibers

Many studies have aimed to improve properties such as strength and modulus of polymer fibers to those of steel in strength and modulus Jiang *et al.* (1993). After the invention of nylon as the first man-made fiber, many fibers, including polyethyleneterephthalate and polyethylene (PE), have been developed and commercialized. Even these generic polymers can, however, achieve high strength levels by drawing. In order to make further progress it is necessary to design the molecular structure and fiber morphology from basic principles. Conditions to fulfill the above mentioned progress are as follows: (i) selecting materials with a molecular backbone linkage that is strong enough (double bond, heterocyclic ring, etc.) and that have a linear conformation and an occupied volume in crystal state that is small and compact; and (ii) drawing the molecules into a fiber such that they are made parallel to the fiber axis, to form a so-called *extended chain structure*.

1. Molecular Design and Morphological Picture

First we illustrate the design of molecule for achieving a high modulus. Relationships between chemical molecular structures and mechanical properties have been discussed since the 1960s (e.g., Tashiro *et al.* (1996) who performed theoretical calculations). In general, a polymer chain can take various forms, such as planar zigzag, helical, linear, and heterocyclic conformations. Theoretical predictions suggest that a strong linkage backbone with linear conformation is one of characteristics for molecules that show high mechanical properties.

Fiber morphological design is discussed next. To explain the relationship between fiber properties and morphology schematically, Hosemann (1962) and Peterlin (1969) derived some models. If a disordered structure exists in the fiber, such as a long-period structure formed by lamella blocks (crystal and amorphous alternating structure), molecular ends, voids, and/or a skin-core structure, they will become sources of strain and stress concentrations once the fiber is deformed. Molecular orientation is one of factors to control the fiber modulus (Northolt and Sikkema 1990). Even chain ends give modulus reduction (Jones and Martin 1995) and actual fiber strength and modulus fail to reach their potential limit values. In the ideal structure of a high-performance fiber, infinitely long molecules make up the crystalline state and are aligned perfectly to the fiber direction.

As examples of fibers that possess the above ideal structure, ultra-high molecular weight PE (UHMW-PE) fibers (Dyneema and Spectra), poly-p-phenyleneterephthalamide (p-aramid) fiber, (Kevlar, Twaron), polypyridobisimidazole (PIDP) fibers, and poly-p-phenylenebenzobisoxazole (PBO) fibers (Zylon) have been developed and commercialized. Their properties are summarized in Table 1 and their repeat units are shown in Fig. 1. It is possible to estimate how much a developed fiber approaches the ideal structure by comparing its mechanical properties with the theoretical values. Theoretical and measured crystal modulus values have been obtained and accumulated systematically on many polymers. Various methods such as lattice mechanics (Tashiro *et al.* 1978, Tashiro and Kobayashi 1996) and a method solving a quantum mechanical equation with approximation Shoemaker *et al.* (1992) are used for this purpose. The crystal modulus can be measured by detecting a change of diffraction spots of fibers under deformation as a function of the applied stress

Table 1
Mechanical properties of typical high-performance fibers.

Fibers	PE	p-Aramid	PBO	Technora	Vectran	PET	Steel
Fiber strength (GPa)	3.0	3.5	5.8	3.1	2.9	1.2	2.8
Fiber modulus (GPa)	172	186	280	70	69	12	200
Measured crystal modulus (GPa)	235	156	480	91	126	95	—
Theoretical modulus (GPa)	295	182	460	—	—	108	—

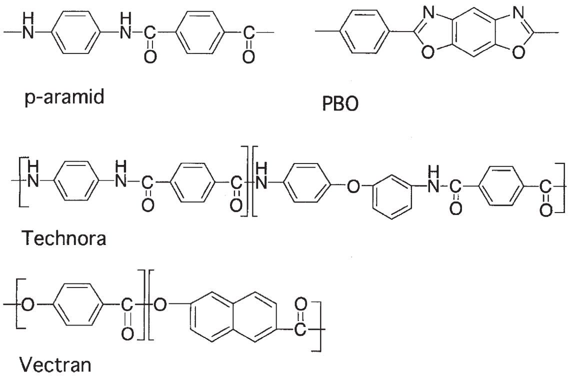


Figure 1
Molecular repeat units.

(Sakurada *et al.* 1962). It has been confirmed that the fiber modulus of p-aramid realizes its theoretical value, while there is still a difference between theory (higher) and experimental values for UHMW-PE and PBO fibers. There are some ideas to explain this difference in terms of structural defects such as molecular misorientation and structural inhomogeneity along the fiber axis, which will be discussed later.

2. *Production*

In order to make molecules aligned to the fiber axis two well-known methods have been developed. In the gel-spinning method flexible chain molecules such as UHMW-PE are spun and drawn with reduction of entanglement densities. The liquid-crystal spinning method makes use of the rigid-rod molecular nature, such as in p-aramid and PBO, that shows liquid crystal state (domains with an alignment of molecules) spontaneously in a certain thermodynamic condition of solution.

2.1 Gel-spinning Method

The development of the gel-spinning method dates back to the invention of ultra-drawing of gel-like fibers by Smith and Lemstra (see *Processing Flexible Polymers to High Performance Fibers**). The fiber is spun from a UHMW-PE solution with a dry-wet, or airgap spinning technology, followed by a drawing process. One of key points of this method is that UHMW-PE (with a molecular weight in excess of one million) is dissolved in a solvent (e.g., xylene) and a quasi-dilute solution is prepared. Being different from the melt state of UHMW-PE, the molecules in the quasi-dilute solution are isolated relative to each other, while only some entanglement between the molecules is present. A gel fiber is then made from this solution by cooling. Molecules in the fiber take on a swollen network structure and the density of the entanglement network is kept low. In addition, the interaction force between PE molecules is so low that an increase in drawing stress of the gel fiber is suppressed even when orientation-induced crystallization occurs. Due to this the gel fiber can be drawn up to by more than 30-fold and the process is called *ultra-drawing*. After drawing the fiber shows high molecular orientation in x-ray diffraction and an absence of a long period in small-angle x-ray scattering (SAXS, Ohta *et al.* 1988) which is characteristic to a melt spun PE fibers.

2.2 Liquid-crystal Spinning Method

As the molecular conformation becomes more rigid, the molecule behaves more like a rod than a coil when in solution; therefore, the excluded volume increases. The thermodynamic state of the solution becomes unstable beyond a certain concentration, and the liquid-crystal domains in which molecules are aligned in a certain direction starts to coexist with the isotropic solution, spontaneously (see *Liquid Crystalline Polymers: An Introduction*$^{\#}$). Molecules in p-aramid and PBO fibers can be extended by utilizing the above feature. The fibers are spun from strong acid solutions with a dry-jet/wet (air gap) spinning method. After extrusion of the spinning solution from a spinneret, the fiber is drawn in the air gap (between the spinneret and the coagulation bath) in which the molecular orientation is increased by the extension flow. Subsequently the drawn molecules are fixed in a coagulation bath with a nonsolvent, followed by a washing process to wash out residual solvent and nonsolvent in the fiber, and dried. The rigid-rod molecules have such a long relaxation time that the extended structure obtained in the air gap can be retained without molecular relaxation. Being different from UHMW-PE fibers, these as spun fibers show a high modulus. To raise the fiber modulus even more, fibers are heat-treated under tension in some cases. Better orientation makes the modulus higher (Northolt and Sikkema 1990).

2.3 Other Methods

One of disadvantages in spinning p-aramid and PBO is that strong acids are required to prepare the spinning solutions. This makes the production equipment longer and more complicated (e.g., the extruder needed to obtain the spinning solution and the coagulation baths) than in a melt spinning process. By reducing the molecular rigidity, by copolymerizing with a flexible spacer into the molecular backbone, the above disadvantage can be moderated. Technora is spun from an isotropic (organic) solution. To raise the molecular orientation of the as-spun fiber, it is drawn up to × 10 with a special drawing technique (Ozawa 1987). The resulting fiber shows a high molecular orientation by x-ray diffraction pattern Nakamae *et al.* (1987). Vectran is spun with a melt spinning method (Chung 1986). Because of the semi-flexible nature in the molecular structure, a quasi-liquid crystal state can be realized in the extension flow in the drawing process and the resulting fiber shows a high modulus and strength.

3. The Relationship Between Fiber Morphology and Properties

The influences of molecular misorientation and of inhomogeneity along the fiber axis to the fiber modulus are surveyed in this part. Figure 2 shows an x-ray diffraction pattern from a PBO (Zylon HM+) fiber. Sharp reflection spots exist on the equator, which demonstrate a high molecular orientation along the fiber axis. The Herman's orientation coefficient has been introduced to express the degree of molecular orientation quantitatively, and measured values of more than 0.99 are reported to be common for UHMW-PE (Busing 1990), p-aramid (Northolt and Sikkema 1990) and PBO Kitagawa *et al.* (1998) fibers. Northolt and Sikkema (1990) treated the relationship between molecular orientation and fiber modulus with a simple model. The method can predict the crystal modulus from measured fiber modulus values in a certain range of molecular orientation by removing the orientation-dependent factor by extrapolation. There exists an agreement between the extrapolated crystal modulus based on this estimation and the crystal modulus directly measured by x-ray diffraction in the p-aramid fiber, while PBO fiber shows a difference. It means that there is still a missing factor, other than the orientation, to explain this gap in PBO fibers having an extended structure.

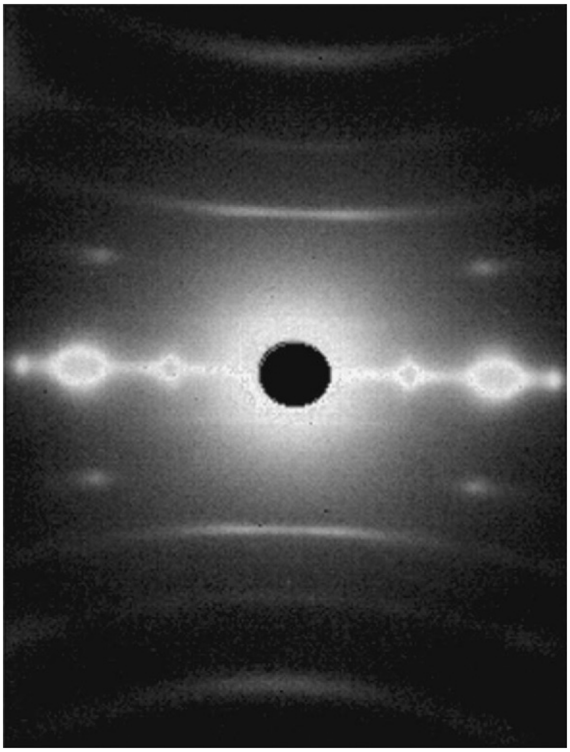


Figure 2
X-ray diffraction pattern of PBO HM+fiber.

Next, structural inhomogeneity along the fiber axis is discussed. It was reported that the intensity of the SAXS pattern has a strong relationship with the presence of a long-period structure and modulus in PE fibers Wu *et al.* (1980). UHMW-PE and p-aramid fibers do not give such SAXS scattering, while PBO HM fiber (heat-treatment under tension) show a four-point SAXS pattern. The PBO molecule is so rigid that there is much less chance to form lamella blocks made of folding molecules, which normally can be detected as the long-period structure. There is still a periodical inhomogeneity along the fiber axis, which gives the SAXS four-point pattern in the PBO case and that also may be the source of hindrance for the fiber modulus to reach the value of its crystal.

As other methods to inspect fibers, such as scanning and transmission electron microscopes are often used. As a method to estimate molecular alignment directly, high-resolution transmission electron microscopy (lattice imaging) has been made available. The fact that p-aramid fiber takes the extended chain structure was first proved by this method Dobb *et al.* (1977). In Fig. 3 a lattice image of a PBO fiber (Zylon HM+) is shown. It can be confirmed that there are fringes parallel to the fiber axis, which reflect the alignment of molecules and persist over 30 nm. It can

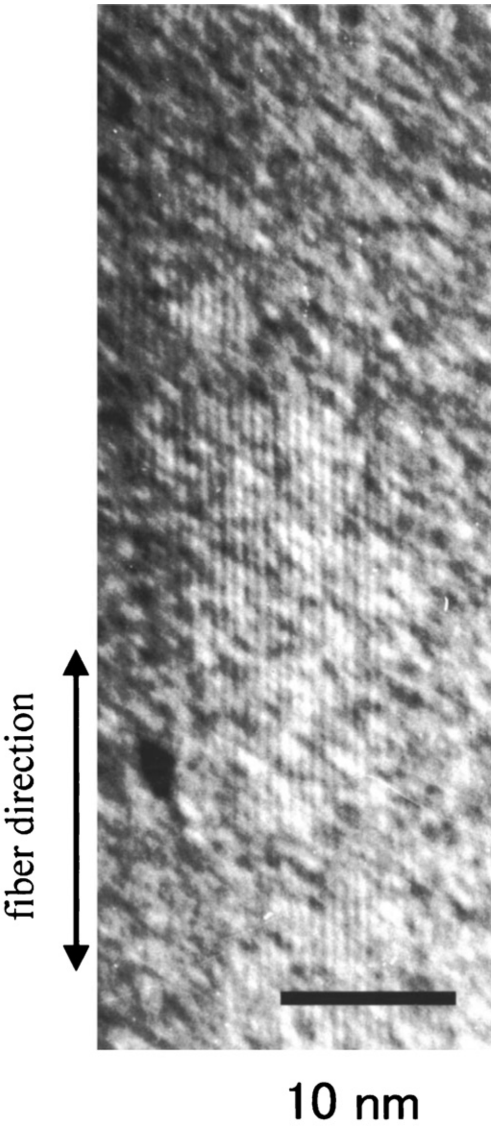


Figure 3
Lattice image of PBO HM + fiber.

be also seen in the figure that some part showing no fringes coexists with crystal regions. This indicates that there still remain defects, such as molecular ends, and conformational disorder. Some models that express the fiber structure features have been proposed and they include all results of inspection of fibers with analytical equipment, for high-performance fibers (Dobb *et al.* 1977, Northolt and Sikkema 1990, Yabuki *et al.* 1975). In the main, attention is paid to the molecular orientation and inhomogeneity along the fiber axis to understand the fiber modulus. Fiber strength is also governed by structural defects. There are some theoretical studies to relate these with the strength Termonia *et al.* (1985). The compressive strength of fibers containing an extended structure is also one of important properties to be improved. It is noteworthy that PIDP fibers give a higher value (1.6 GPa) than any of the other super fibers (~ 0.6 GPa) (Sikkema 1998). It is plausible that hydrogen bonding in PIDP fiber forms interactions between neighboring molecules and reinforces the compressive strength, while p-aramid fiber (also having hydrogen bonding) shows a lower value than PIDP. Further study is needed to explain the above discrepancy quantitatively from the viewpoints of morphology and theory.

4. Concluding Remarks

Selecting a suitable molecule is the first step to start developing a high performance fiber. The fiber should take an extended structure that shows high molecular orientation and homogeneity along the fiber axis. As far as the Young's modulus is concerned, a high crystal modulus is the target to be achieved in developing high performance fibers and it is realized as the actual modulus in some fibers. Existing theories can estimate the theoretical modulus with fair accuracy. But concerning other properties, they need to be further refined to estimate, for example, compressive strength.

See also: Processing Flexible Polymers to High Performance Fibers*; Processing Rigid Polymers to High Performance Fibers*; Solid-state Extrusion to High Performance Fibers$^{\#}$; Main-chain Liquid Crystalline Polymers for Fibers$^{\#}$

Bibliography

Blackwell J, Biswas A 1987 Structure and properties of thermotropic liquid crystalline copolyesters. In: Ward I M (ed.) *Development in Oriented Polymers 2*. Elsevier, London, Chap. 5, pp. 153–97

Busing W R 1990 X-ray diffraction study of disorder in Allied Spectra-1000 polyethylene fibers. *Macromolecules* **23**, 4608–10

Chung T 1986 The recent developments of thermotropic liquid crystalline polymers. *Polym. Eng. Sci.* **26**, 901

Dobb M G, Johnson D J, Saville B P 1977 Supramolecular structure of a high-modulus polyaromatic fiber (Kevlar 49). *J. Polym. Sci. Polym. Phys.* **15**, 2201–11

Flory P J 1953 *Principles of Polymer Chemistry*. Cornell University Press, New York

Hosemann R 1962 Crystallinity in high polymers, especially fibers. *Polymer* **3**, 349–92

Jiang H, Adams W W, Eby R K 1993 High performance polymer fibers. In: Cahn R W, Haasen P, Kramer E J (ed.) *Materials Science and Technology*. VCH, New York, Chap. 13, pp. 596–652

Jones M -C G, Martin D C 1995 Molecular stress and strain in an oriented extended-chain polymer of finite molecular length. *Macromolecules* **28**, 6161–74

Kitagawa T, Ishitobi M, Yabuki K 2000 An analysis of deformation process on poly-p-phenylenebenzobisoxazole (PBO) fiber and a structural study of the new high-modulus type PBO HM+ fiber. *J. Polym. Sci., Polym. Phys.* **38**, 1605

Kitagawa T, Murase H, Yabuki K 1998 Morphological study on poly-p-phenylenebenzobisoxazole (PBO) fiber. *J. Polym. Sci. Polym. Phys.* **36**, 39–48

Lemstra P J, Kirschbaum R 1985 Speciality products based on commodity polymers. *Polymer* **26**, 1372–84

Lemstra P J, van Aerle N A J M, Bastiaansen C W M 1987 Chain-extended polyethylene. *Polymer* **19**, 85–98

Nakamae K, Nishino T, Shimizu Y, Matsumoto T 1987 Experimental determination of the elastic modulus of crystalline regions of some aromatic polyamides, aromatic polyester, and aromatic polyester ketone. *Polymer* **19**, 451–9

Northolt M G, Sikkema D J 1990 Lyotropic main chain liquid crystal polymers. *Adv. Polym. Sci.* **98**, 115–77

Ohta T, Kunugi T, Yabuki K 1988 *Koukyoudo Koudanseiritsu Seni.* Kyoritsu, Tokyo

Ozawa S 1987 A new approach to high modulus, high tenacity fibers. *Polymer J.* **19**, 119

Peterlin A 1969 Bond rupture in highly oriented crystalline polymers. *J. Polymer Sci. Polym. Phys.* **7**, 1151

Sakurada I, Nukushina Y, Ito T 1962 Experimental determination of the elastic modulus of crystalline regions in oriented polymers. *J. Polym. Sci.* **57**, 651–60

Shoemaker J R, Horn T, Haaland P D, Pachter R P, Adams W W 1992 Electronic structure and properties of strained polymers: 1. Methods for polyethylene. *Polymer* **33**, 3351–6

Sikkema D J 1998 Design, synthesis and properties of a novel rigid polymer, PIDP or 'M5': high modulus and tenacity fibers with substantial compressive strength. *Polymer* **39**, 5981–6

Tadokoro H 1979 *Structure of Polymers.* Wiley-Interscience, New York

Tashiro K, Kobayashi M 1996 Molecular theoretical study of the intimate relationship between structure and mechanical properties of polymer crystals. *Polymer* **37**, 1775–86

Tashiro K, Kobayashi M, Tadokoro H 1978 Calculation of three-dimensional elastic constants of polymer crystals. *Macromolecules* **11**, 908–20

Termonia Y, Meakin P, Smith P 1985 Theoretical study of the influence of the molecular weight on the maximum tensile strength of polymer fibers. *Macromolecules* **18**, 2246–52

Wu W, Simpson P G, Black W B 1980 Morphology and tensile property relations of high-strength/high-modulus polyethylene fiber. *J. Polym Sci. Polym. Phys.* **18**, 751

Yabuki K, Ito H, Ota T 1975 Studies on the fine structure and the physical properties of poly (p-phenylenetrephthalamide) (Part 1). *Sen-i Gakkaishi* **11**, 20

T. Kitagawa
Toyobo Research Center, Shiga, Japan

Polymer Fibers: Formation and Structure

Conventional fiber formation starting from semicrystalline polymers consists of three processes: melt-spinning, drawing, and annealing. The drawing and annealing processes are usually combined, as shown schematically in Fig. 1. The spinning and the drawing and annealing processes also can be conducted continuously, and are then known as a spin-draw process.

In the spinning process, either melt- or solution spinning, polymers are extruded through tiny orifices to form thin threads and taken up using a winder. Since polymer molecules are not well aligned along the fiber axis at this stage, as-spun fibers usually exhibit poor mechanical properties such as low modulus, low tenacity, and high elongation. Therefore, the drawing and annealing processes are necessary to fabricate fibers with well-developed structure.

In the drawing and annealing processes, as-spun fibers pass through several sets of rolls with different surface velocities. Fiber temperature is controlled either by heated rolls, plate heaters, or some other heating device. Drawing brings about molecular orientation and mechanical anisotropy, whereas annealing enhances the crystallization and thermal stability. Both properties are needed as fundamental characteristics of the finished fibers.

If the take-up velocity is increased significantly in the melt spinning process, the extensional stress applied to the spinning line increases and the as-spun fibers start to have high molecular orientation. At the same time, since the crystallization rate increases with an increase in molecular orientation, crystallization can be promoted in the spinning line even though the cooling of the molten polymer is extremely fast. This type of crystallization is defined as orientation-induced crystallization. Because of significant fiber structure development, a high-speed melt-spinning process can provide fibers directly applicable for their end use (Ziabicki and Kawai 1985).

1. *History of the Fiber Structure Model*

Fiber structure models were initially developed for natural fibers. A model proposed for cellulose fibers was the "fringed micelle model". The most important structural difference between synthetic and natural fibers is in the existence or nonexistence of periodic structure along the fiber axis (Kaji 1997).

The "fringed-micelle micro-fibril model," however, cannot explain (i) the SAXS intensity on the equator, and (ii) the mechanical properties of fibers. The "paracrystalline layer structure model," in which the concept of folded-chain crystal is introduced, was for the crystalline region.

"Shish-kebab structures," in which extended chains form the "shish," and crystalline lamellae epitaxially grown from the side surface of the shish form the "kebab," have become widely accepted. This structure was originally called the "row structure," as all the shish are oriented along the flow direction. Considering the overlapping of lamellae grown from adjacent shish, the "shish-kebab model"

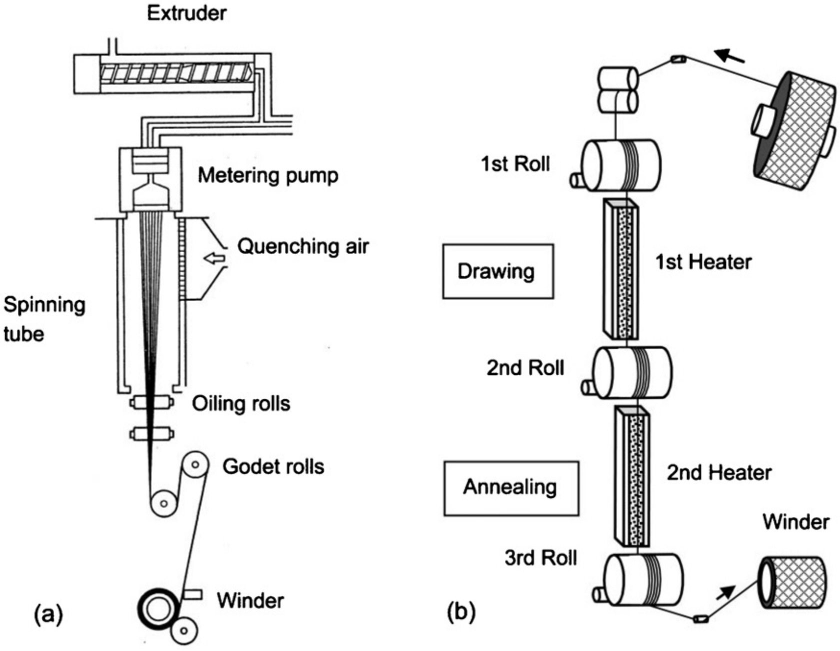


Figure 1
Schematic illustrations of (a) melt spinning and (b) drawing and annealing processes.

is also referred to as the ''interlocking shish-kebab model'' (Bashir *et al.* 1986).

Although the shish-kebab structure has been widely observed in highly oriented polyethylene (PE), polypropylene (PP), and polyamide 66 (PA66), further investigation may be necessary concerning the applicability of this model for polymers with lower crystallizability, such as aromatic polyesters.

2. *Aromatic Polyesters*

2.1 Fiber Processing and Structure Development

Melt-spinning processes can be applied to the production of fibers from aromatic polyesters such as poly(ethylene terephthalate) (PET), poly(trimethylene terephthalate) (PTT), poly(butylene terephthalate) (PBT), and poly(ethylene naphthalate) (PEN). These polymers have relatively low crystallization rates. The glass transition temperatures of these polymers are higher than room temperature except for PBT, which has a glass transition temperature close to room temperature. Therefore, the ordinary melt-spinning process of aromatic polyesters provides amorphous fibers with low molecular orientation. Development of the fiber structure can be achieved by the drawing of as-spun fibers at a temperature near the glass transition temperature, followed by the annealing of the drawn fibers at a temperature between the glass transition and melting temperatures. A temperature close to the maximum-crystallization-rate temperature may be preferable for the annealing process.

In high-speed melt spinning of these polymers, molecular orientation increases with an increase in the take-up velocity. Changes in the birefringence of various aromatic polyester fibers with an increase in the take-up velocity are summarized in Fig. 2. The difference in birefringence development behavior originates from the differences in molecular rigidity, intrinsic birefringence, etc. In these polymers, orientation-induced crystallization starts to occur at the take-up velocities of around 2–5 km min^{-1}. In each polymer, the starting point of crystallization varies slightly, depending on the extrusion and cooling conditions and the molecular weight of the polymers. After the onset of orientation-induced crystallization, a higher take-up velocity generally yields higher molecular orientation, higher crystallinity, and higher ordering in the crystalline region. Conversely, molecular orientation in the amorphous region is relatively low in comparison with that in the fibers produced through conventional drawing and annealing processes.

As a typical example, structural changes in PET fibers from the high-speed melt-spinning process are summarized in Fig. 3. Initiation of orientation-induced crystallization is represented by the change in the WAXS pattern and a sudden increase of density at around 4 km min^{-1}. Birefringence of as-spun PET

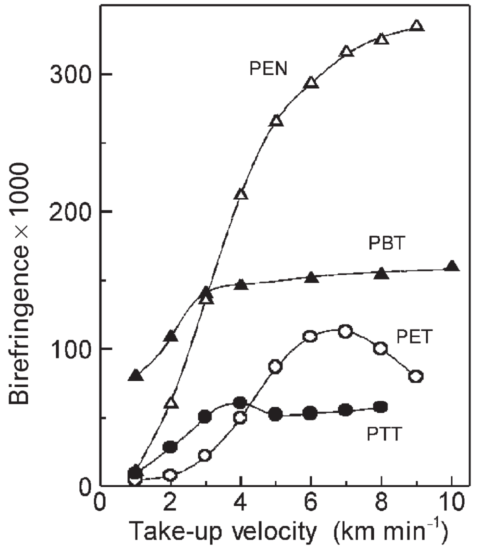


Figure 2
Birefringence development of various aromatic polyester fibers with increasing take-up velocity in the high-speed melt spinning process.

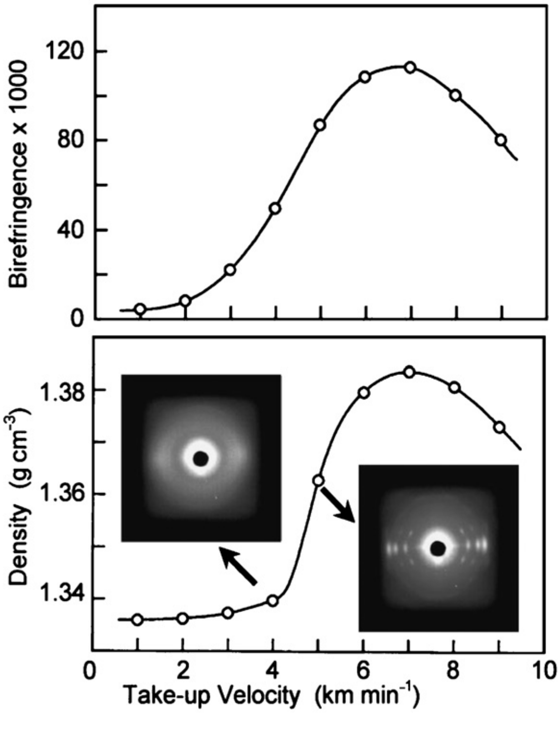


Figure 3
Change in fiber structure of PET in high-speed melt spinning process.

fibers is roughly predictable from the stress–optical coefficient and the tensile stress acting on the spinning line near the solidification point. When the take-up velocity exceeds 7 km min^{-1}, however, molecular orientation, crystallinity, and mechanical properties start to decrease. The development of significant structural variation was found in the cross-section of PET fibers prepared in this take-up velocity region.

2.2 Crystal Modification

The crystal unit cell of all the aromatic polyesters discussed above is triclinic. No polymorphism is reported for PET and PTT, while PBT and PEN are known to have α- and β-form crystals. In comparison with PET, PTT has an extremely low crystalline modulus in the chain direction. This is because the aliphatic part of the crystalline structure of PET has a trans conformation, whereas that of PTT has a highly coiled structure of gauche–gauche conformation (Desborough 1979). Therefore, the unit cell parameters of PTT show a distinct mechanical response upon the application of tensile stress to the fibers (Jakeways 1975).

PBT fibers prepared both through the conventional process and the high-speed melt-spinning process have α-form structure (Fujimoto *et al.* 1988). The crystal form changes from α to β when a tensile strain of about 12% is applied, but reverts to α form upon the removal of strain (Yokouchi *et al.* 1976). As the c-axis length of the β-form unit cell is about 8% longer than that of the α form, the stress–strain curves of PBT fibers exhibit a remarkable yielding behavior showing a distinct plateau region. The yielding stress corresponds to the critical transition stress from α to β form (Tashiro *et al.* 1980).

In the case of PEN, α-form crystals are formed in the conventional spinning and drawing processes, whereas the coexistence of α and β form crystals is confirmed in the high-speed spun fibers. Even though the β-form unit cell of PEN has a shorter c-axis length and higher crystalline density than the α form, the amount of β-form crystals increases with an increase in the take-up velocity, that is, with an increase in the tensile stress applied during the orientation-induced crystallization. In the quiescent state, β-form crystals can be developed only when the crystallization temperature is higher than 220–245 °C, whereas the formation of the β form occurs at much lower temperatures in the high-speed spinning line (Miyata *et al.* 1997).

3. Tilting

Another interesting feature in the structure of aromatic polyester fibers is the tilting of the c-axes of the crystalline unit cell away from the fiber axis. This means that the molecular chains on average are not

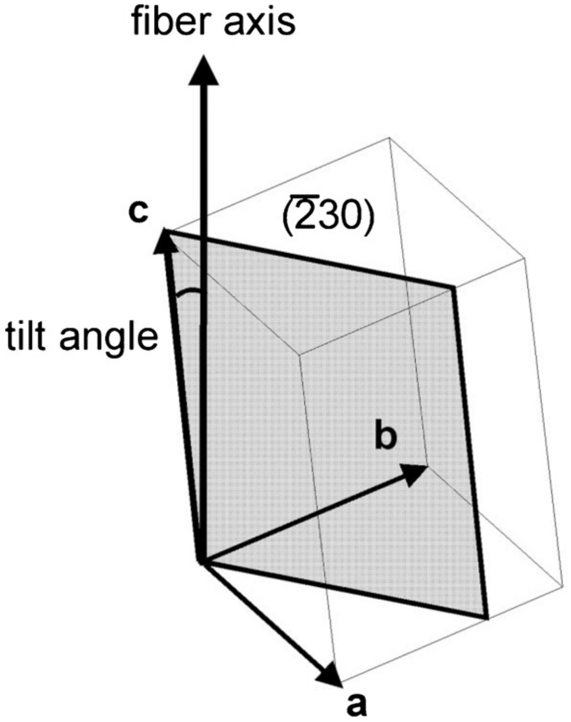


Figure 4
Schematic illustration explaining the tilting of c-axis along $(\bar{2}30)$ plane in PET fibers.

aligned parallel to the fiber axis. The direction and degree of tilting can be analyzed from the shifts of crystalline reflections from layer lines in the WAXS pattern. In PET fibers, tilting of the c-axis occurs along the $(\bar{2}30)$ plane as shown schematically in Fig. 4. The tilt-angle of conventional PET fibers is about 5°, whereas that of high-speed spun fibers is reported to decrease from 10° to 2° with an increase in the take-up velocity (Shimizu *et al.* 1984). In PBT fibers, tilting of about one degree was found both in α- and β-form crystals. The direction of tilting is along the $(\bar{2}10)$ and (130) planes, respectively. The tilt-angle of high-speed spun PBT fibers also decreases with an increase in the take-up velocity (Fujimoto *et al.* 1988). In PET and α-form PBT fibers, tilting is essentially parallel to the crystallographic plane containing the benzene ring. On the other hand, tilting of about 3–5° occurs along the $(\bar{1}20)$ plane in PTT fibers, although the molecular chain in the crystal is in 2_1 helix conformation.

4. Aliphatic Polyesters

Some of the aliphatic polyesters, which generally have lower glass transition and melting temperatures in comparison with aromatic polyesters, attract interest because of their biodegradability. Melt spinning of these polymers, however, is difficult to control, as the extruded fibers tend to reach the take-up device in a rubbery state. This is caused by a combination of two factors: the low crystallization rate, and the glass transition temperature being lower than the room temperature. Crystallization of these polymers may be controlled through the installation of a hot tube or hot water-bath into the spinning line.

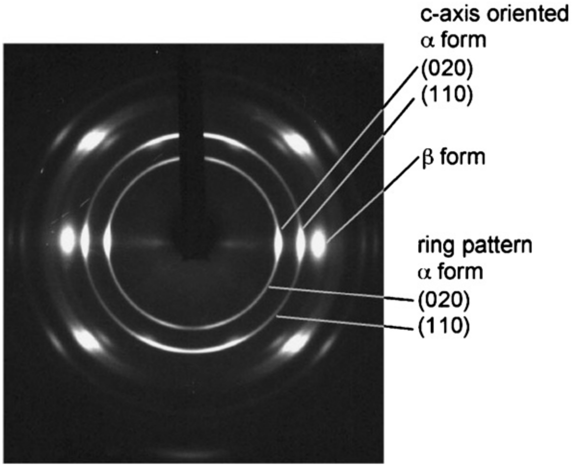


Figure 5
Bimodal orientation of α form crystals of PHB in drawn and annealed P(HB/HV) fibers. Coexistence of β form crystals is also indicated.

4.1 Poly(3-hydroxybutylate)

Poly(3-hydroxybutyrate) (PHB), and copolymers of PHB with 3-hydroxyvalerate (P(HB/HV)), are produced by certain bacteria. The fiber structure of PHB and P(HB/HV) is mainly composed of α-form crystals of PHB, but small amounts of β-form crystals are known to coexist. The molecular chain in the α-form crystal is in 3_1 helix conformation, whereas that in the β-form crystal is almost in the planar zigzag conformation (Orts *et al.* 1990). The α- and β-form crystals usually exhibit ordinary c-axis orientation, and the amount of β form varies depending on the drawing and annealing conditions (Furuhashi *et al.* 1998) or the take-up velocity in the high-speed spinning process (Schmack *et al.* 1999).

In some cases, the α-form crystals in drawn and annealed fibers exhibit peculiar bimodal orientation, which is represented by the coexistence of diffraction spots and rings in the WAXS pattern. Some rings are complete, with a higher intensity near the meridian, while other rings are incomplete, with a lack of intensity near the meridian, as shown in Fig. 5. Detailed analyses of the ring pattern reveal that the c-axis is preferentially oriented perpendicular to the fiber axis, whereas the a- and b-axes are randomly distributed around the c-axis (Furuhashi *et al.* 1998).

4.2 *Polylactide*

Poly(L-lactide) (PLLA) has a glass transition temperature higher than room temperature. PLLA fibers prepared through conventional fiber formation processes and high-speed melt-spinning processes are known to exhibit α-form crystals, whereas the β-form crystals can be developed in the fibers drawn to extremely high draw ratios (Hoggsteen *et al.* 1990). The molecular chain in the α- and β-form crystals is in 10_3 and 3_1 helix conformation, respectively.

5. *Aliphatic Polyamides*

5.1 *Fiber Processing and Structure Development*

Fiber formation behavior in aliphatic polyamides (PA) is sensitive to the ambient conditions. This is due to the highly hygroscopic nature of the polymers. Structural parameters and mechanical properties are also sensitive to humidity because of the change in water content in the fibers.

In ordinary melt-spinning processes with polyamide 6 (PA6) and polyamide 66 (PA66), amorphous as-spun fibers can be obtained, as the glass transition temperature of these polymers is around 40 °C, which is higher than room temperature. (In the case of PA66, which has a significantly higher crystallization rate than PA6 in the quiescent condition, spherulitic structures may be developed in the spinning line if the spinning conditions result in a relatively low cooling rate.) On exposure to moisture after spinning, however, diffusion of water into the amorphous as-spun fibers results in crystallization. This process is referred to as the conditioning process.

When as-spun fibers of PA6 and PA66 are prepared at take-up velocities from 1 to $3\,\mathrm{km\,min^{-1}}$, birefringence increases significantly and fibers elongate spontaneously during the conditioning process. The amount of spontaneous elongation may be suppressed through the enhancement of crystallization in the spinning process, by applying water to the surface of running fibers before winding. Drawing of the as-spun fibers of PA6 and PA66 usually is conducted at room temperature. Therefore, the drawing process is called cold drawing. The reduction of the glass transition temperature by moisture uptake and heat generation during the drawing process provides enough mobility to molecules for the formation of crystalline fiber structure. Higher annealing temperatures are necessary for the production of high-strength polyamide fibers, and highly oriented and crystallized fibers can be obtained in a similar manner to the aromatic polyesters.

5.2 *Crystal Modification*

PA66 has crystal modifications of α and β forms. Both the drawn and annealed fibers and high-speed spun fibers (Shimizu *et al.* 1985) exhibit α-form structure. Very weak (002) reflections of the β form were found to coexist in the WAXS pattern of high-speed spun fibers.

Various crystal modifications such as α, β, γ, and smectic (pseudo-hexagonal) forms have been reported for PA6. A reasonable rationalization is that there are only two basic crystallographic forms, α and γ. Although there are some differences in reflection intensities, the unit cell parameters of α and β forms are identical. Although γ and smectic form crystals have different thermal stabilities, these two forms exhibit similar WAXS patterns. In drawn and annealed PA6 fibers, α and β forms coexist, and the proportion of α form increases with increases in the crystallization temperature and applied tension.

In high-speed melt-spinning processes, orientation-induced crystallization starts to occur at take-up velocities of around 3–$4\,\mathrm{km\,min^{-1}}$. In this region, the γ form dominates the crystalline structure of as-spun fibers, and the proportion of α form increases with an increase in take-up velocity (Shimizu *et al.* 1985). Higher molecular weight also yields greater proportions of α-form crystals (Koyama *et al.* 1986). These results suggest that the higher tensile stress and higher crystallization temperature lead to the preferred formation of the α form. Heuvel *et al.* (1976) confirmed that orientation distribution of PA6 α-form crystals is axis-symmetric with respect to the fiber axis, and there is no predominant direction in the rotation of the unit cell.

6. *Polyolefins*

Polyolefins such as polyethylene (PE) and polypropylene (PP) have high crystallization rates and their glass transition temperatures are below room temperature. Therefore, these polymers need to be crystallized in the spinning line to obtain solid fibers. Although the usage of PE fibers is limited, PP fibers are widely used as nonwoven fabrics and other applications such as carpets. In the spun-bond process, which is one of the important technologies in producing nonwoven fabrics, melt-spun fibers are directly spread on a wire mesh to obtain fabrics.

Although a wide range of drawing temperatures can be applied for the drawing of as-spun PE and PP fibers, drawing temperatures near the crystalline dispersion temperature may be preferable because the starting fibers already have high crystallinity.

6.1 Polyethylene

When the tensile stress applied to the spinning line is low or the take-up velocity is low, the crystallographic b-axis tends to orient perpendicular to the fiber axis, whereas the a-axis tends to orient along the fiber axis. The orthorhombic unit cell rotates around the b-axis gradually with an increase in the tensile stress or the take-up velocity, and eventually ordinary c-axis oriented crystals can be formed, as shown in Fig. 6. It should be noted that the density or crystallinity of PE fibers hardly changes, irrespective of the extensive variation in fiber formation conditions. Drawn PE fibers usually exhibit high molecular orientation with high c-axis orientation of the crystals.

Recent improvements in polymer synthesis technologies have brought high spinnability to PE polymers. Structure development in PE high-speed spinning processes varies slightly depending on the type of PE polymers: high density PE, linear-low- density PE, or metallocene-catalyzed linear low-density PE. In the case of linear low-density PE, formation of a-axis oriented crystals in the low speed region was not detectable (Cho *et al.* 2000).

6.2 Polypropylene

PP generally has a lower crystallization rate than PE. More importantly, the crystallization rate of PP is comparable to the cooling rate the polymer experiences in the spinning line. Therefore, the development of crystalline structure in the melt-spinning process depends strongly on the spinning conditions and polymer characteristics.

When the take-up velocity is low, and therefore the cooling rate and the applied tensile stress are low, monoclinic crystals can be formed in the as-spun fibers. The crystalline orientation is nearly isotropic in these fibers. Formation of monoclinic crystals suggests that the crystallization proceeds at relatively high temperatures. When the take-up velocity is increased, the crystalline structure changes to the smectic (pseudo-hexagonal) form, because of the enhancement of cooling and a reduction in the crystallization temperature. There can be additional development of smectic crystals even after the spinning process, because the glass transition temperature of PP is below room temperature. Further increase of the take-up velocity leads to a significant development of molecular orientation. The molecular orientation enhances the crystallization, and because of the increase in the crystallization temperature, the crystal structure changes from isotropic smectic crystals to highly oriented monoclinic crystals (Cao *et al.* 1989).

The full series of structural changes described above can be observed when the molecular weight is low, and the extrusion temperature and throughput rate are high. Under these conditions, the mechanism of orientation-induced crystallization becomes dominant only at high take-up velocities. If molecular weight is high, the extrusion temperature and throughput rate are low; then the crystalline form simply remains monoclinic, and the crystalline

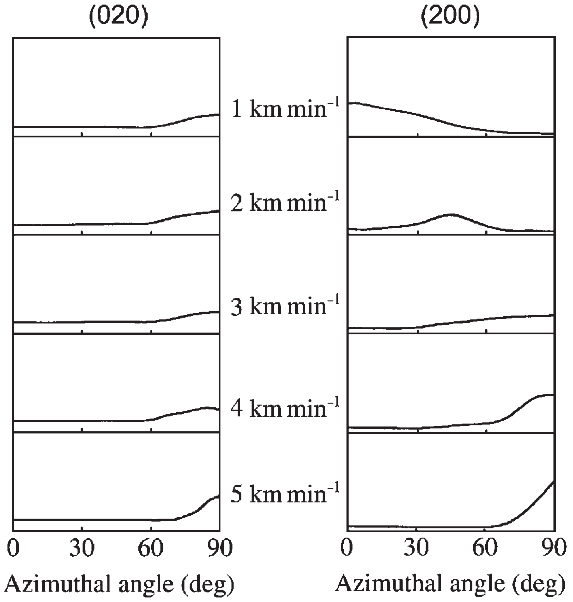


Figure 6
Changes in azimuthal intensity distribution of (020) and (200) reflections of PE fibers in wide-angle x-ray diffraction measurements. Take-up velocities are indicated.

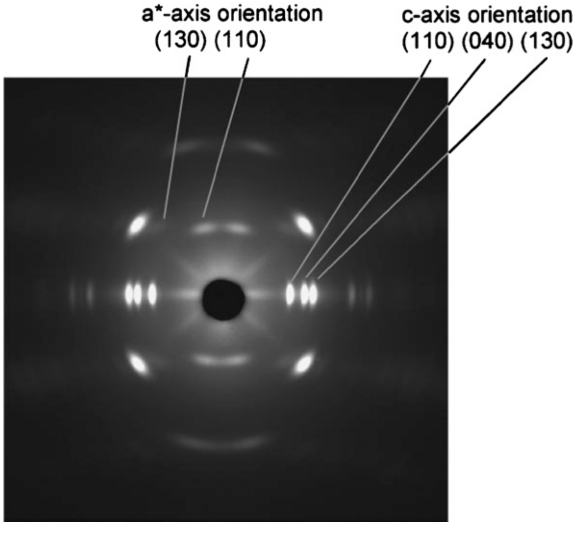


Figure 7
Bimodal orientation in PP fibers.

orientation changes from nearly isotropic to a highly oriented state with an increase in the take-up velocity.

Another important feature of PP fibers is the development of bimodal orientation, i.e., the coexistence of c- and a*-axis oriented monoclinic crystals. Bimodal orientation can be confirmed from WAXS patterns in which the (110) reflection splits to two peaks, one on the equator and the other near the meridian, whereas (040) reflection remains on the equator as shown in Fig. 7. This structure resembles the so-called crosshatch structure, in which epitaxial growth of daughter lamellae occurs on the side-surface of oriented parent lamellae.

On-line measurements of the development of crystalline structure were conducted utilizing a synchrotron radiation source (Kolb *et al.* 2000). The experimental results suggested that the development of c- and a*-axis-oriented crystals proceeds almost simultaneously. The amount of a*-axis-oriented

Table 1
Crystal modifications appearing in fiber structure.

		Unit-cell parameters			
Polymer	Crystal modification crystal system	a (nm) α	b (nm) β	c (nm) γ	L_r (nm) ρc (g cm^{-3})
Aromatic polyesters					
PET	#	0.456	0.594	1.075	1.075
	Triclinic	98.5	118	112	1.455
PTT	#	0.46	0.62	1.83	1.83
	Triclinic	98	90	112	1.43
PBT	α (*)	0.483	0.594	1.159	1.159
	Triclinic	99.7	115.2	110.8	1.404
	β	0.495	0.567	1.295	1.295
	Triclinic	101.7	121.8	99.9	1.283
PEN	α	0.651	0.575	1.32	1.32
	Triclinic	81.3	144	100	1.407
	β (*)	0.926	1.559	1.273	1.273
	Triclinic	121.6	95.6	122.5	1.439
Aliphatic polyesters					
PHB	α (*)	0.576	1.32	0.596	0.278
	Orthorhombic	90	90	90	1.26
	β	0.522	0.522	#	#
	Hexagonal	90	90	120	#
PLA	α (*)	1.06	0.61	2.88	0.288
	Orthorhombic	90	90	90	1.2845
	β	1.03	1.82	0.9	0.3
	Orthorhombic	90	90	90	1.286
Aliphatic polyamides					
PA66	α (*)	0.49	0.54	1.72	1.72
	Triclinic	48.5	77	63.5	1.24
	β	0.49	0.8	1.72	1.72
	Triclinic	90	77	67	1.25
PA6	α (*)	0.956	1.724	0.801	1.724
	Triclinic	90	67.5	90	1.23
	γ	4.79	4.79	1.67	1.67
	Hexagonal	90	90	60	1.17
Polyolefins					
PE	#	0.74	0.493	0.253	0.253
	Orthorhombic	90	90	90	1
PP	α (*)	0.665	2.096	0.65	0.65
	Monoclinic	90	99.3	90	0.938
	γ	0.65	0.65	#	#
	Hexagonal	90	90	120	#

L_r: chain length /repeating unit.
*: Preferred crystalline modification in high-speed melt spinning.

crystals varies depending on spinning conditions, whereas the highly drawn and annealed fibers usually exhibit only the c-axis-oriented crystals.

Although the molecular orientation of as-spun fibers strongly depends on the tensile stress applied to the spinning line, birefringence of as-spun fibers is significantly larger than that predicted from the applied stress and stress-optical coefficient of the polymer melt (Shimizu *et al.* 1985). The nonvalidity of the stress-optical rule suggests that crystallization is accompanied by a significant development of molecular orientation.

7. Concluding Remarks

Fiber structure in semicrystalline polymers has been explained, based on the degree of molecular orientation and differences in crystalline structure. Crystal modifications appearing in the fiber structure are summarized in Table 1. It is still difficult to find a general rule between the unit cell and the type of crystal modifications appearing in the fiber processing. It appears, however, that the crystal form with higher density is preferred in high-speed melt-spinning, especially in the extremely high-speed region. The only exception in Table 1 is PA66. It should be noted here that the unit cells with longer fiber repeat distances are not necessarily preferred, although the crystallization is strongly accelerated by high tensile stress.

There are more complicated and hierarchical aspects in fiber structure. Important features that could not be included in the previous explanations are the concepts of entanglement and network structure.

Long and Ward (1991) found that the true-stress vs. true-strain curves of PET fibers prepared under a wide range of processing conditions can be superposed by simply shifting the curves along the true-strain axis, as shown in Fig. 8. The entanglement density estimated from the network draw ratio, which can be estimated from the amount of shift for curve matching, and thermal shrinkage stress are roughly comparable to those obtained from the plateau modulus measured through the rheological measurements of the polymer melt. These results indicate that PET fibers retain a level of high-order structure that is not affected by processing.

Fundamental studies of orientation-induced crystallization behavior suggest that the starting of crystallization is not controlled simply by the degree of molecular orientation, but that the nonlinear characteristics of elongational viscosity and the state of molecular entanglement also play an important role (Okamoto *et al.* 1998). In this context, the "shish" in the "shish-kebab" structure corresponds to the tightened molecules between two adjacent entanglements, that is, the starting point of the crystallization. Therefore, the role of entanglement should be considered more extensively for the better understanding of the mechanism of fiber structure development. Several types of mechanical measurements have been proposed for the analysis of entangled structures. However, at present there are no structural measurement techniques applicable for the direct analysis of the state of entanglement.

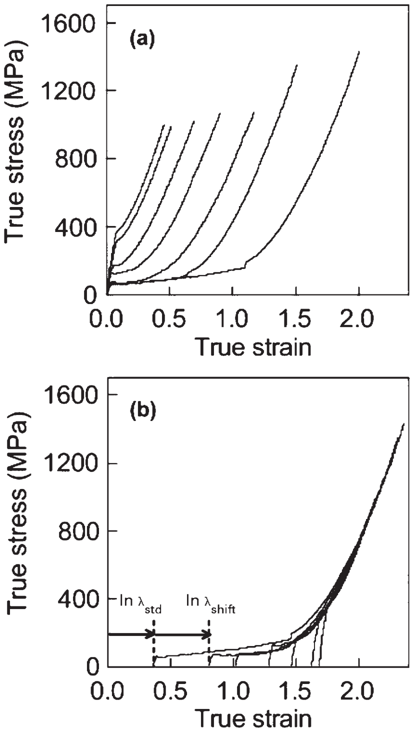


Figure 8
(a) True stress versus true strain curves of PET fibers prepared by various processing conditions. (b) Matching of true stress versus true strain curves of PET fibers. Superposed curve corresponds to the network draw ratio of unity. $\lambda_{std,}$, network draw ratio of standard fibers; $\lambda_{shift,}$, draw ratio corresponding to axial shift along the true strain axis needed for curve matching.

Bibliography

Bashir Z, Hill M J, Keller A 1986 Comparative study of etching techniques for electron microscopy using melt processed polyethylene. *J. Mater. Sci. Lett.* **5**, 876–8

Cao J, Kikutani T, Takaku A, Shimizu J 1989 Nonisothermal orientation-induced crystallization in melt spinning of polypropylene. *J. Appl. Polym. Sci.* **37**, 2683–97

Cho H H, Kim K H, Ito H, Kikutani T 2000 Fine structure and properties of polyethylene fibers in high-speed spinning. *J. Appl. Polym. Sci.* **77**, 1195–206

Desborough I J, Hall I H, Neisser J Z 1979 The structure of poly(trimethylene terephthalate). *Polymer* **20**, 545–52

Fujimoto K, Iohara K, Owaki S, Murase Y 1988 Structure and properties in high-speed melt spinning of poly(butylene terephthalate). *Sen'i Gakkaishi* **44**, 394–401

Furuhashi Y, Ito H, Kikutani T, Yamamoto T, Kimizu M, Cakmak M 1998 Structural analysis of poly(3-hydroxybutyrate-co-3-hydroxyvalerate) fibers prepared by drawing and annealing processes. *J. Polym. Sci.* **36**, 2471–82

Heuvel H M, Huisman R, Lind K C J B 1976 Quantitative information from x-ray diffraction of nylon-6 yarns. *J. Polym. Sci., Polym. Phys. Ed.* **14**, 921–54

Hoggsteen W, Postema A R, Pennings A J, Brinke G T 1990 Crystal structure, conformation, and morphology of solution-spun poly(L-lactide) fibers. *Macromolecules* **23**, 634–42

Jakeways R, Ward I M, Wilding M A, Hall I H, Desborough I J, Pass M G 1975 Crystal deformation in aromatic polyesters. *J. Polym. Sci.* **13**, 799–813

Kaji K 1997 From the fringed-micelle microfibril model to the shish-kebab model. *Sen'i Gakkaishi* **53**, P132–6

Kolb R, Seifert S, Strbeck N, Zachmann H G 2000 Simultaneous measurements of small-and wide-angle x-ray scattering during low speed spinning of poly(propylene) using synchrotron radiation. *Polymer* **41**, 1497–505

Koyama K, Suryadevara J, Spuriell J E 1986 Effect of molecular weight on high-speed melt spinning of nylon-6. *J. Appl. Polym. Sci.* **31**, 2203–29

Long S D, Ward I M 1991 Tensile drawing behaviour of polyethylene terephthalate. *J. Appl. Polym. Sci.* **42**, 1911–20

Mezghani K, Spruiell J E 1998 High-speed melt spinning of PLLA filaments. *J. Polym. Sci. Pt. B—Polym. Phys.* **36**, 1005–12

Miyata K, Kikutani T, Okui N 1997 Fiber structure formation in ultra-high-speed melt spinning of poly(ethylene-2,6-naphthalene dicarboxylate). *J. Appl. Polym. Sci.* **65**, 1415–27

Okamoto M, Kubo H, Kotaka T 1998 Elongational flow-induced crystallization of poly(ethylene terephthalate) under supercooled state. *Polymer* **39**, 3135–42

Orts W J, Marchessault R H, Bluhm T L, Hamer G K 1990 Observation of strain-induced β form in poly(β-hydroxyalkanoates). *Macromolecules* **23**, 5368–70

Schmack G, Tändler B, Vogel R, Beyreuther R, Jacobsen S, Fritz H G 1999 Biodegradable fibers of poly(L-lactide) produced by high-speed melt spinning and spin drawing. *J. Appl. Polym. Sci.* **73**, 2785–92

Shimizu J, Kikutani T, Takaku A, Okui N 1984 Fiber structure formation in high-speed melt spinning of poly(ethylene terephthalate): characterization of fiber structure by considering oriented mesophase. *Sen'i Gakkaishi* **40**, T177–84

Shimizu J, Okui N, Kikutani T 1985 Fine structure and physical properties of fibers melt-spun at high speeds from various polymers. In: Ziabicki A, Kawai H (eds.) *High-Speed Fiber Spinning*. Wiley, New York, pp. 429–83

Tashiro K, Nakai Y, Kobayashi M, Tadokoro H 1980 Solid-state transition of poly(butylene terephthalate) induced by mechanical deformation. *Macromolecules* **13**, 137–45

Yokouchi M, Sakakibara Y, Chatani Y, Tadokoro H, Tanaka T, Yoda K 1976 Structure of two crystalline forms of poly(butylene terephthalate) and reversible transition between them by mechanical deformation. *Macromolecules* **9**, 266–73

Ziabicki A, Kawai H (eds.) 1985 *High-Speed Fiber Spinning*. Wiley, New York

T. Kikutani
Tokyo Institute of Technology, Japan

Polymer Matrix Composites with Roughly Equiaxed Reinforcements: Microstructure and Mechanical Behavior

Mineral particulate fillers are finely divided solids (with an aspect ratio of about unity) which are added to a polymer matrix in relatively large amounts to modify its properties and/or to reduce the cost of the resulting compound. Generally, a filler material is stiffer than the matrix and depending on its nature, shape, and treatment, it is reinforcing or not. In the latter case one speaks of extenders. Fillers or extenders are primarily used with three goals in mind, namely to improve dimensional stability (heat distortion temperature, shrinkage, thermal expansion), mechanical properties (sample stiffness, compressive strength, and—if possible—toughness), and/or the optical or electrical properties of the composite material. The most frequently used isomorphic fillers are calcium carbonate, silicates, and glass spheres. For comparison, occasionally data will be cited concerning kaolin, talc, or hydrous alumina, which have a flake-like shape.

The addition of particulate fillers influences all stages of the fabrication, and the structure and performance of the resulting composites. Since, as mentioned above, fillers and extenders are generally stiffer than the polymer to which they are added, the focus here is on the effects of a stiff second phase on the matrix structure and on deformation, creep, and failure mechanisms of the composite. Thus, elastomeric modifiers, polymer blends, and filled elastomers are not considered; these are discussed elsewhere. Only a brief reference will be made to stress analysis of heterogeneous composites and to the various approaches to estimate their elastic properties, since these are treated elsewhere (*Composites, Physical Properties of**).

A concise description of particulate mineral fillers is given in many publications (e.g., Katz and Milewski 1978, Schlumpf 1993, Rothon 1995, *Mineral Fillers in Thermoplastics* 1999, Eurofillers 1999, Pukanszky 1995). The main filler characteristics to be considered when used in a polymer matrix are the following:

- chemical composition (mineralogical origin and purity, possible catalytic actions);
- shape, often characterized by the aspect ratio (spherical, irregular, fibrous, plate-like);
- mean particle diameter, size distribution, largest particles (top cutoff diameter);
- hardness, density; and
- particle surface properties (surface energy, specific surface area, surface treatment through adhesion promoters or release agents, type of bonding that can be engaged with the matrix, affinity to humidity or stabilizing agents, and also matrix nucleation efficiency).

When optimizing the above parameters attention must be paid to synergistic effects, and also to indirect and frequently nonlinear effects on polymer structure and mobility. The main effects resulting from the addition of (particulate) fillers comprise:

- an influence on processing (compounding, melt or resin viscosity, abrasive action, sedimentation and dispersion behavior, change of reactivity and heat capacity, dimensional stability, and industrial safety problems like dust);
- structure modification (effects on morphology, density, and optical and electrical properties);
- small strain behavior (stress distribution, elastic moduli, stiffness);
- deformation and rupture (creep, interfacial debonding, micro shear yielding, crack pinning); and
- physical and environmental aging (humidity effects, chemical degradation).

The microstructure of a particulate-filled polymer (see Fig. 1) is characterized by two main parameters: the filler volume fraction, $\Phi = V_f/V_c$, where the subscripts f and c refer to filler and composite, respectively, and V is volume; and the average particle diameter, d, or the average interparticle distance, D (these two parameters are interrelated for a given Φ_f and filler particle distribution, and the latter parameter is also quantifiable, see *Composites, Microstructure of: Quantitative Description**). A microstructural feature more difficult to quantify is the possible alteration of the intrinsic matrix microstructure that may result from the presence of filler particles. One example of such an alteration is the formation of an interlayer of matrix material having a higher density or a specific morphology through epitaxial growth on the filler surface.

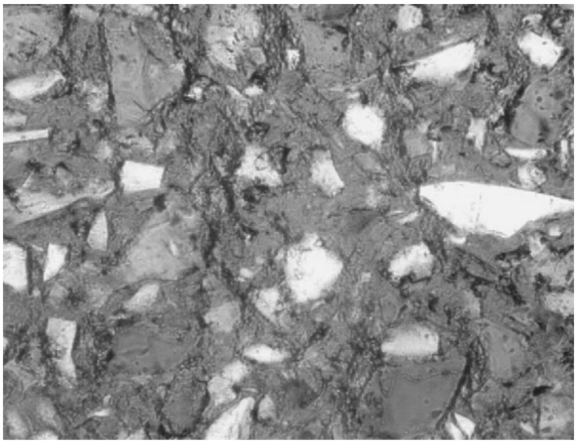


Figure 1
Typical microstructure of a polymer composite filled with irregularly shaped particles (polished surface of an epoxy resin filled by 40 vol.% of quartz powder with an average particle size of 50 μm, loaded in tensile creep at 105 °C).

Both the nature of the filler and the microstructural parameters exert a significant influence on the mechanical properties of the resulting polymer matrix composite, ranging from small strain elastic properties (see *Composites, Physical Properties of** and Rothon 1995, 'Mineral Fillers' 1999, Eurofillers 1999, Pukanszky 1995, Nielsen and Landel 1994) to the ultimate properties at large deformation.

1. Elastic and Viscoelastic Properties and Matrix Morphology

1.1 Influence of the Filler on Elastic and Viscoelastic Properties

To a first approximation, the effects of inert, noninteracting filler particles on scalar properties of a composite can be correlated with the volume fraction, Φ_f, of the filler particles (see *Composites, Physical Properties of**). Thus the density, ρ, and the heat capacity, c_p, of a composite are calculated in a straightforward manner from the rule of mixtures (attention must, however, be paid to the possible inclusion of air or humidity into the compound during mixing).

The static elastic moduli (shear, Young's, or bulk) of a polymer filled with rigid particles increase with the filler volume fraction, Φ_f. The effect of apparently inert, isomorphic particles is well represented by the modified Kerner equation (see Nielsen and Landel 1994):

$$E_c = E_1 \frac{1 + AB\Phi}{1 - B\Psi} \quad \text{with} \quad A = \frac{7 - 5v}{8 - 10v} \quad \text{and}$$

$$B = \frac{(E_f/E_m) - 1}{(E_f/E_m) + A} \tag{1}$$

where

$$\Psi = 1 - \exp\left(\frac{-\Phi_f}{1 - \Phi_f/\Phi_{max}}\right)$$

and the subscripts c, m, and f refer to the composite, the matrix, and the filler respectively, v is Poisson's ratio of the matrix, and Φ_{max} is the maximum packing fraction of the filler particles in the given situation. This modified Kerner equation has been shown to be valid for a range thermoplastic and thermosetting matrices (Nielsen and Landel 1994, Vollenberg and Heikens 1989, Stricker *et al.* 1997). However, there are also a number of alternative approaches for the calculation of composite moduli (see *Composites, Physical Properties of**).

The combined effects of time and the presence of fillers on the viscoelastic behavior are frequently considered separately. Thus, the time-dependent compliance, $D_c(t)$, of a particulate composite is to a first approximation proportional to the matrix creep deformation, $\varepsilon_m(t)$, and the ratio of the (instantaneous) elastic moduli, E_m/E_c (Nielsen and Landel 1994):

$$\sigma_0 D_c(t) = \varepsilon_c(t) = \varepsilon_m(t) E_m / E_c \qquad (2)$$

The relaxation moduli, $G(t)$ and $E(t)$, can be represented in a similar way. In dynamic experiments, however, it is frequently found that the damping peak is broadened (or that new peaks arise) and that the phase lag of internal friction, defined by the term tan δ, increases nonlinearly with extension owing to strain concentrations at and between particles, and decohesion at particle–matrix interfaces. Thus, Stricker *et al.* (1997) have observed in sPP filled with glass beads a peak some 50 K above the glass transition temperature of 6 °C that is purely related to interfacial loss. Dewetting also occurs during creep, contributing a strain component, ε_d, which is superimposed on the viscoelastic strains, $\varepsilon_m(t)$, of the matrix material. Notwithstanding the different origins of these two contributions, the resulting creep compliances agree with the prediction of the time–temperature–stress superposition (TTSS) principle (see Sect. 3.2).

1.2 Filler Effects on Thermoplastic Matrix Microstructure

Filler particles provide a very large (specific) surface area, and their interaction with the surrounding matrix gives rise to different interfacial phenomena (cf. *Composites: Interfaces*#). Of particular importance in the present context are wetting and adhesion and effects on polymer organization and morphology. These effects can be amplified and, to a certain extent, controlled by adequate filler surface treatment.

Several colloidal metals and salts as well as active fillers such as talc and aluminum oxide act as nucleating agents in crystallizable polymer melts (Galeski 1995, Pukanszky *et al.* 1994). The most frequently used extender, $CaCO_3$, has relatively low nucleation efficiency (Pukanszky *et al.* 1994). The important feature of a good nucleating agent seems to be the existence of alternating rows of polar and nonpolar groups at the seed surface (Galeski 1995). The main effects resulting from the presence and homogeneous distribution of a large number of crystallization nuclei are a higher crystallization temperature and a much finer spherulitic structure. The overall crystallinity, x_c, is influenced in different ways. Glass beads in sPP (Stricker *et al.* 1997) lead to a decrease of x_c since the total interspherulitic or impingement area (which has a lower degree of crystallinity) is increased. $CaCO_3$ in PP (Pukanszky *et al.* 1994) has very little influence on x_c and talc-filled composites show a diverging behavior (Stricker *et al.* 1997, Pukanszky *et al.* 1994). Evidently for any given filler, nucleation effects should be proportional to the total interface area. However, there seems to be a lower limit to particle size. Very small particles in a coarse spherulitic material are observed to congregate in the interspherulitic regions giving rise to brittle interspherulitic fracture (Friedrich 1978).

The morphology of the matrix material in contact with a structured surface may be influenced by epitaxial or transcrystalline growth. Heteroepitaxial growth is well known to occur at polymer–polymer interfaces (e.g., PE–iPP (Petermann 1995, Yan *et al.* 1998)). Here the criterion for secondary nucleation to occur is that the substrate crystal dimension in the matching direction must be greater than the crystal thickness of the depositing polymer. Once this requirement is satisfied, subsequent epitaxial growth is based on the lamellar growth habit of the depositing polymer itself (Petermann 1995, Yan *et al.* 1998). Epitaxial phenomena have also been found in a number of mineral-filled semicrystalline thermoplastics (HDPE, PP, and PA filled with $CaCO_3$ or SiO_2 particles, see Sect. 5.1 and Galeski 1995, Michler 1992, Bartczak *et al.* 1999, Walter *et al.* 1997). Filler surface modification is also used for the selective formation of a given crystal phase structure in replacement of another. Thus the β-crystallization of polypropylene is favored by the presence of surface-treated calcium carbonate (a mixture of equal amounts of pimelic acid and calcium stearate acts as a β-nucleator for PP) (Tjong *et al.* 1997, Labour 1999). Whereas the above elements of sample morphology generally have a modest effect on the small strain properties, they influence more strongly the deformation mechanisms and thus the yield behavior and toughness of the filled polymer (see Sects. 4.2 and 5).

The presence of an interface in amorphous as well as in semicrystalline polymers frequently causes local modifications which are more difficult to detect than nucleation or epitaxial growth. These interfacial effects concern the reduced segment mobility, improved local packing, and modified forces of attachment. Thus the strong effect of filler size and composition on Young's modulus of polystyrene composites is ascribed to segmental orientation and improved packing in the neighborhood of the filler surfaces (Vollenberg and Heikens 1989). An increase of density of an interfacial matrix region (by some 10%) has an exponential effect on the stiffness of that region—and thus of the composite. The interlayer (or interphase) concept is essentially based on a micromechanical analysis of filler effects (Michler 1992, Bartczak *et al.* 1999, Walter *et al.* 1997, Tjong *et al.* 1997, Labour 1999, Jancar *et al.* 1992, Pukanszky 1993, Vörös *et al.* 1997). Typically interphase thicknesses of 0.1 μm (PP) to 0.5 μm (PVC) have been assigned (Vörös *et al.* 1997). Large specific surface areas help to increase the elastic moduli and reduce creep rates. High free energy surfaces (metals, mineral fillers) induce crystallinity and provide stronger physical bonds. However, for the same reason these may give rise to unwanted adsorption (of water or additives such as stabilizers or cross-linking agents). Depending on their chemical composition, fillers may catalyze secondary or degradative reactions that influence the molecular structure and stability of the surrounding matrix (Katz and Milewski 1978, Eurofillers 1999, Joppien 1978). However, finely ground $CaCO_3$ may also adsorb and deactivate remnants of catalysts. Therefore, much effort has been concentrated on filler surface modification by impregnation, grafting, or addition of a (elastomeric) compound susceptible to form an interphase around the particles (Katz and Milewski 1978, Schlumpf 1993, Rothon 1995, 'Mineral Fillers' 1999, Eurofillers 1999, Pukanszky 1995, Nielsen and Landel 1994, Schlumpf 1994, Michler 1992, Jancar *et al.* 1992, Pukanszky 1993, Vörös *et al.* 1997). These techniques result in better control of filler–matrix adhesion and interfacial activity and an improvement in processing conditions (particle dispersion, rheology) and final performance.

2. *Creep and Damage Mechanisms*

2.1 Basic Filler-related Deformation and Damage Mechanisms

Small strain deformation of a filled composite is essentially borne by the matrix. With increasing strain, the strongly heterogeneous deformation will give rise to local damage at the site of the largest stress concentration. The basic damage mechanisms observed are particle–matrix debonding (often termed "dewetting" even though the matrix is a solid), void formation and coalescence (see Fig. 2(a)), and plastic deformation of the matrix material (Fig. 2(b)). The essential intrinsic parameters governing these mechanisms and the stress level at which they occur are particle size and distribution, interfacial adhesion, and matrix deformability. Small, well-dispersed and weakly adhering particles initiate a large number of small and stable voids, which permit further deformation of the matrix ligaments between particles (Fig. 2(b)). Voids created from large particles in a brittle matrix may give rise to a critical crack and to instantaneous brittle fracture (Fig. 2(c)). This is one of the (evident) reasons why the top cutoff value of the particle size distribution should be well controlled. Similar problems are created by aggregates and by matrix-free interfacial regions at high degrees of filling (Fig. 2(d)).

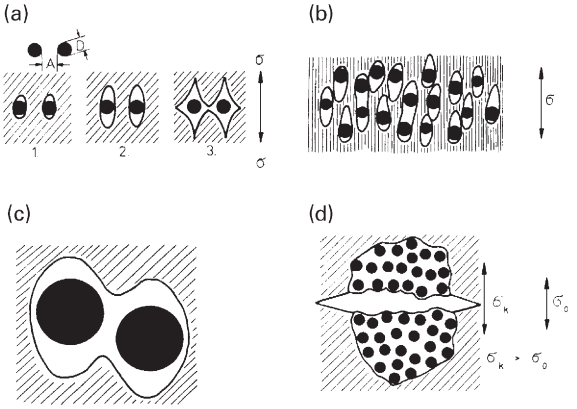


Figure 2
Schematic representation of void formation from filler particles of different sizes: (a) particle–matrix debonding, matrix void formation, and coalescence; (b) fibrillar deformation of interparticular ligaments; (c) critical voids form around large particles or (d) agglomerates (Michler 1992).

Interfacial void formation in composites with well-bonded fillers requires larger strains. The stability of the strained composite then depends on particle size and ligament properties. Again, large particles are susceptible to act as defects giving rise to brittle fracture—although at a much higher stress than the non-adhering particles. Small, well-bonded particles may delay void formation to such a stress level that the first debonding event gives rise to brittle fracture. Thus the formation of voids that are well distributed in space and time contributes to toughness (see Sect. 5). The above general remarks are valid for both thermoplastics and thermosets; characteristic differences only become apparent when discussing fracture phenomena.

2.2 Creep Behavior

At short times and small loads, creep is dominated by the viscoelastic strain, $\varepsilon_m(t)$, of the matrix material (Eqn. (2)). This means that at identical stresses the creep rate of the composite is smaller than that of the pure matrix by a factor of E_m/E_c. The response of the composite is linear up to a critical threshold strain, ε_{th}, which will vary with the system and the external parameters σ and T (Smith 1989, Kaiser *et al.* 1989, Gregory *et al.* 1999, Vogel *et al.* 1994). (Vogel *et al.* (1994) used a cyclic loading method to determine the incurred damage from the loss of stiffness. For a glass mat-reinforced vinyl ester resin no damage was observed below a threshold strain, ε_{th}, which at room temperature amounted to 0.48%. For the same material, damage has also been assessed by plotting isochronal creep strains (Gregory *et al.* 1999); the deviation from linearity starts at the same value of 0.48%.) The threshold strain generally marks the onset of interfacial debonding, which contributes a strain component ε_d. The interfacial debonding strain, ε_d, is superimposed on the viscoelastic creep deformation increasing the nonlinearity of the response. As an example the creep behavior of a quartz flour-filled epoxy resin (CY 225 from

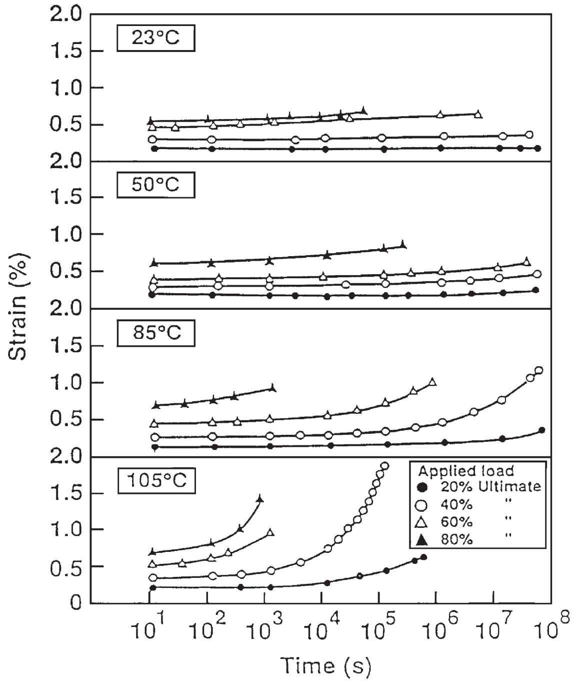


Figure 3
Tensile creep response of a (42 vol.%) quartz flour-filled epoxy resin (temperature range 23–105 °C and for stresses between 20% and 80% of the respective short-time strengths) (Smith 1989, Kaiser *et al.* 1994).

Ciba-Geigy, a difunctional hydantoin-based epoxy resin cured with anhydride hardener Hy 925 filled with varying degrees of quartz flour (W10 from Quarzwerke Frechen, Germany)) is considered (Fig. 3).

Notwithstanding the different origins of the two contributions to creep, such creep curves can still often be reduced by TTSS to a single creep compliance master curve (Smith 1989, Kaiser *et al.* 1989). (A successful reduction of creep data has also been achieved for a vinyl ester resin charged with glass fiber mat (Gregory *et al.* 1999).)

3. *Effects of Filler on "Strength"*

The term "strength" designates the engineering stress at failure. Depending on the nature of the matrix material and the filler parameters (degree of filling, particle size, bonding), failure occurs in a brittle or in a ductile manner. Consequently, one has to distinguish between brittle strength, σ_{cb}, the strength, σ_y, at the yield point (in tension or compression), and the ultimate strength, σ_{ub}, after appreciable plastic deformation and/or stable crack growth. The latter phenomena and filler effects on toughness are discussed in Sect. 5.

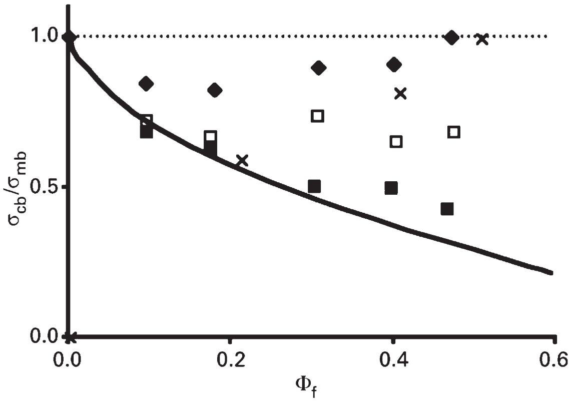


Figure 4
Tensile strengths of differently filled epoxy resins: ×, 50 μm silica particles (Moloney *et al.* 1987); and 62 μm glass beads: ♦, treated with A 187; ■, treated with release agent; □, untreated (Spanoudakis and Young 1984).

3.1 *Brittle Fracture*

In tension, most particulate-filled cross-linked resins and rigid thermoplastics break in a brittle manner before reaching the yield point. Brittle fracture is generally triggered by the appearance and/or activation of a critical defect. Several criteria have been proposed to estimate the brittle strength of such composites.

(*a*) *Critical matrix strain, ε_{mb}, in the stress direction.* Provided there is good adhesion between a rigid matrix and filler particles, it can be assumed (Katz and Milewski 1987, Pukanszky 1995, Michler 1992, Jancar *et al.* 1992) that the composite breaks when the average matrix ligament in the stress direction attains its ultimate failure strain, ε_{mb}. The overall composite strain, ε_{cb}, at this moment is

$$\varepsilon_{cb} = \varepsilon_{mb}\left(1 - \sqrt[3]{\frac{3}{\pi}}\,\Phi_f^{1/3}\right) \tag{3}$$

The composite breaking stress, σ_{cb}, then equals $\varepsilon \cdot E_c$. In most—but not all—composites σ_{cb} decreases with increasing filler contents and average particle size (Fig. 4) (Schlumpf 1994, Spanoudakis and Young 1984, Moloney *et al.* 1987, Cantwell *et al.* 1988, Nakamura and Yamaguchi 1993, Karger Kocsis 1995, Kinloch and Young 1983). The above approach suffers from the fact that the ligaments which break first are not average ligaments, that ligament fracture is in competition with interfacial debonding and microshear banding (see below), and that the composite can most probably tolerate several interfacial debonding events. Much more than Young's modulus, composite strength is sensitive to the quality of adhesive bonding, to particle size and distribution, to the presence of defects (inherent flaws), and to matrix properties (see below).

(*b*) *Inherent flaws.* Using a linear elastic fracture mechanics approach, brittle fracture of any material containing a defect of size *a* can be related to the toughness, K_{Ic}, of the material through

$$\sigma_{\mathrm{cb}} = \frac{K_{\mathrm{Ic}}}{Y\sqrt{a^{1/2}}} \tag{4}$$

The limited strength of a homogeneous material is then explained through the existence of inherent flaws, of which the size, *a*, has been related to that of pre-existing heterogeneities or defects such as the particle diameter, poorly bonded or weak particles, or flaws generated during sample preparation. Unfortunately, the resulting values of *a* turn out to be systematically too large; for example, for epoxy resins filled with particles between 5 μm and 50 μm, the values of *a* range between 36 μm and 522 μm (Moloney *et al.* 1987, Kinloch and Young 1983).

Closer approximations of elastic properties and strength can be obtained in several ways, by statistical mechanical analysis (Nairn 1999), finite element analysis, and numerical calculations (Pukanszky 1995, Jancar *et al.* 1992, van Mier *et al.* 1993) (see also *Composites, Physical Properties of**). Such approaches also consider internal stresses resulting from the thermal mismatch of fillers and matrices. In any event, failure prediction requires knowledge of the relevant failure mechanisms and of adequate failure criteria.

3.2 *Yield Stress and Ductile Failure*

There exist several models to predict the yield stress of particulate composites (Pukanszky 1995, Nielsen and Landel 1994, Stricker *et al.* 1997, Pukanszky *et al.* 1994, Jancar *et al.* 1992, Pukanszky 1993, Vörös *et al.* 1997, Joppien 1978, Nicolais and Narkis 1971, Liang and Ly 2000). Of these, the modified approach originally proposed by Nicolais and Narkis, the composition-dependent model of Pukanszky 1993 and Vörös *et al.*, and the finite element calculations of Jancar *et al.* will be discussed here.

Any homogeneous or filled (polymer) specimen yields if and when local shear deformation initiates and extends sufficiently far before crazing and/or fracture dominate the deformation behavior. In the simplest model it is considered that the stresses are transmitted by the matrix and by inclusions in proportion to their effective cross-sections. In tension, the composite yields (at a stress σ_{yt}l) when the average matrix stress reaches the matrix yield stress σ_{ym}. Stress concentration (caused by unbonded inclusions or aggregates) are taken into account through a factor $S<1$ and the quality of bonding and possible effects on the surrounding matrix material by the stress component σ^* transmitted through the particles (Pukanszky 1995, Stricker *et al.* 1997, Jancar *et al.* 1992, Vörös *et al.* 1997, Nicolais and Narkis 1971, Liang and Ly 2000):

$$\sigma_{\mathrm{yt}} = \sigma_{\mathrm{ym}}(1 - 1.21\Phi_{\mathrm{f}}^{2/3})S + \sigma^* 1.21\Phi_{\mathrm{f}}^{2/3} \tag{5}$$

For unbonded particles, σ^* is zero. The remaining term can be considered as a lower bound of σ_{yt} (Nicolais and Narkis 1971). Syndiotactic polypropylene (sPP) of low stereoregularity filled with 5 μm glass beads (Stricker *et al.* 1997) and LDPE and PVC filled with 3.6 μm $CaCO_3$ filler particles (Vörös *et al.* 1997) comply quite well to the upper bound of the above equation (Pukanszky 1995, Stricker *et al.* 1997, Vörös *et al.* 1997).

Considering the composition dependence of tensile yield stress including crystallinity, crystal size, possible filler anisotropy, and the quality of adhesion, Pukanszky *et al.* (1994) arrived at the following relationship:

$$\sigma_{\mathrm{yt}} = \sigma_{\mathrm{ym}} \frac{1 - \Phi_{\mathrm{f}}}{1 + 2.5\Phi_{\mathrm{f}}} \exp(B\Phi_{\mathrm{f}}) \tag{6}$$

where *B* is an empirical parameter characterizing the degree of particle–matrix interaction (it increases with the specific surface area and the adhesive strength of the interphase).

As stated above, macroscopic yielding requires that the local yield zones grow, overlap (between neighboring particles), and finally extend into a larger volume. Sjoerdsma (1989) and Jancar *et al.* (1992) used finite element simulation to analyze the stress distribution within a regular array of particles. From the octahedral shear yield criterion they then determined the shape of the yielded zone around each particle as a function of external stress. Knowing the interparticle distance they obtained the overlap distance, *c*, and the overlap *area*, *F*(*c*), within each cell. From their analysis they derived the following equation for a yielding composite:

$$\sigma_{\mathrm{yt}} = \sigma^*_{\mathrm{ym}}[1 + 0.33F(c)\Phi_{\mathrm{f}}^2] \tag{7}$$

where σ^*_{ym} is the yield strength of the matrix material containing a few isolated filler particles. There are three restrictions to the validity of Eqn. (7): the quality of adhesion between matrix and fillers, the competition with crazing and brittle fracture (especially at moderate and high filling fractions), and the existence of an upper limit of yield stress if σ_{yt} approaches $1.33\sigma^*_{ym}$ (Jancar *et al.* 1992). Data from experiments at 80 °C with maleated PP filled with $CaCO_3$ (Jancar *et al.* 1992) comply well with Eqn. (7).

3.3 Ultimate Strength

The effect of fillers on the ultimate properties of a thermoplastic composite is situated between two extremes. Nonadhering particles in a ductile matrix have little influence on *ultimate* tensile fracture behavior. Thus plasticized PVC shows no change in ε_{cb} (about 365%) and tensile strength σ_{cb} (19.5 MPa) up to 35% by weight of $CaCO_3$ (Nielsen and Landel 1994). Strong improvements in strength (especially of PE and PP) owing to strain hardening of the matrix material can be obtained with suitably chosen particulate charges. Introducing a strain hardening parameter, λ^n, into Eqn. (7), Pukanszky (1990) obtained in such cases the following empirical equation for tensile strength:

$$\sigma_{ub} = \lambda^n \sigma_{ym} \frac{1 - \Phi_f}{1 + 2.5\Phi_f} \exp(B\Phi_f) \qquad (8)$$

3.4 Compressive Strength

Failure in compression occurs generally in shear. For modest filling fractions and a ductile polymer matrix one can expect the compressive strength to increase with the filling fraction and to be superior to the yield strength, σ_{yt}, in tension. This is owing to two effects: the increase in matrix shear strength owing to hydrostatic compressive stresses and the increase in shear viscosity owing to the presence of filler particles. The difference between tension and compression is even larger with brittle matrices, which in tension are sensitive to crazing and/or to brittle fracture owing to internal flaws or interfacial debonding effects. Thus for 40% glass beads in PA 66, strength values in tension of 98 MPa and in compression of 252 MPa have been reported (Nielsen and Landel 1994) (the compressive strength of the pure matrix is 29 MPa). The corresponding data for a 40% silica-filled epoxy resin are 64 MPa in tension, 154 MPa in compression, and 98 MPa for the pure resin (Moloney *et al.* 1987).

4. Filler Effects on Crack Propagation and Toughness

The term toughness of a composite is understood to mean its resistance to crack extension. Toughness is traditionally expressed by either of the fracture mechanics quantities K_{Ic} or G_{Ic}. It should be noted, however, that G_{Ic} is much more sensitive to particle–matrix adhesion than K_{Ic}. It is necessary, therefore, to discuss the mechanisms leading to crack extension. In view of their differences it is appropriate to discuss crack extension in thermoplastics and in cross-linked resins separately.

4.1 Filled Thermoplastics

The main deformation mechanisms of an unfilled semicrystalline thermoplastic subjected to dilatational strains are elastic dilatation—and eventually cavitation—of the amorphous regions and concomitantly straining of the crystal lamellas. The material will be tough if, at this stage, crystal plastic deformation (fine shear, block shear, decrystallization) and subsequent formation of fibrillar structures dominate with respect to localized crazing and/or brittle fracture (Michler 1992, Kinloch and Young 1983, Kim and Michler 1998a, 1998b, Kausch *et al.* 1999). The essential point in filling a semicrystalline thermoplastic to improve toughness is to favor microplastic deformation at a high stress level. This requires, first, good adhesion of the matrix to the filler particles so as to attain high stress levels and avoid premature crack extension through void coalescence (see Fig. 2(b)). Good adhesion can be obtained from adhesion promoters, surface treatment, or by increasing the amorphous component of the polymer—which often adheres more easily to mineral substrates (Katz and Milewski 1978, Schlumpf 1993, 1994, Rothon 1995, 'Mineral Fillers' 1999, Eurofillers 1999, Pukanszky 1993, 1995, Stricker *et al.* 1997). Secondly, particle size and volume fraction should be controlled to favor ductile deformation of interparticle ligaments.

Generally, the toughest material is obtained with incompletely monodispersed medium-sized particles (but the definition of medium sized has to be discussed). At low rates of deformation where local yielding takes place, K_{Ic} or G_{Ic} usually increase up to filling fractions of 15% (Pukanszky 1995, Jancar *et al.* 1992, Fekete *et al.* 1999).

In impact loading, however, toughness tends to decrease strongly with increasing Young's modulus (as shown by α- and β-PP (Stricker *et al.* 1997, Labour 1999)). A number of HDPE compounds give a more surprising response. Studies (Bartczak *et al.* 1999) have shown that both toughness and Young's modulus of a calcium carbonate-filled HDPE can be increased *simultaneously*. A similar enhancement of toughness by small, well-bonded particles has been found (Walter *et al.* 1997). These observations have been explained by the particle-induced formation of an interphase favoring extensive matrix plastic deformation (Michler 1992, Bartczak *et al.* 1999, Walter *et al.* 1997, Kim and Michler 1998a, 1998b).

4.2 Filled Thermosets

The network structure of cross-linked polymers is characterized by the higher density of cross-linking points (as compared to the density of entanglement points in thermoplastics) and the consequent reduced deformability of the chain segments. As indicated in Sect. 3.1, in a strained filled thermoset essentially two local deformation mechanisms are observed: particle–matrix debonding and microshear banding.

Basic fracture mechanisms can be well observed in creep at constant load. Fracture starts from a pre-existing defect (frequently at the surface and not always visible). The defect grows (often rather slowly) forming a subcritical crack. The tip of this crack constitutes a process zone, which—as in a pure resin (Moloney *et al.* 1987, Kinloch and Young 1983, Cantwell and Roulin-Moloney 1989)—is subject to stress concentration and local plastic deformation leading to crack tip blunting. In addition it can be concluded from fractographic and microscopic analyses (Smith 1989, Cantwell *et al.* 1988) that debonding occurs ahead of the crack tip. During this phase of growth the crack extends from one debond to the other more or less within a plane. The resulting fracture surface of such a debonded zone is rather smooth. This phase of crack growth is thermally activated, strongly time dependent, and generally consumes most of the time to fracture of the specimen. In a constant load experiment, the stress intensity factor, K_c, and energy release rate, G_c, will grow with crack length, a. At a certain critical crack length a_{II}, G_c becomes larger than the material resistance and the excess released energy is used to drive the crack even faster: the crack therefore accelerates (Cantwell *et al.* 1988). This acceleration marks the transition from the second to the third phase of crack growth. The associated crack length, a_{II}, can be used to calculate K_{Ic} from Eqn. (4). In fact, there is a rather good correlation between the stress intensity factor, $K_c(a)$, and the rate of crack propagation, da/dt, which can be used to evaluate the expected lifetime of a loaded sample by integration over da up to a_{II} (Beer *et al.* 1986).

The above conclusions are confirmed by an analysis of the stress-induced damage in the bulk material. Depending on the volume damage ahead of the crack, on the toughness of the matrix, and on temperature, there are three modes of crack extension, which are illustrated in Fig. 5. For the same quartz flour-filled epoxy resin discussed in Sect. 2.2, it has been concluded (Smith 1989, Smith *et al.* 1988) that at low temperatures, high stresses, and short times to fracture a single debonding event can initiate rapid brittle fracture (Fig. 5(a)). The main sample volume then remains essentially undamaged. At intermediate temperatures a limited number of debonding events occur before the onset of crack propagation (in this case there is limited volume damage). The crack then jumps from one debond to another, more or less within a plane and a smooth subcritical crack zone is formed (Fig. 5(b)). The third fracture mode occurs at high temperatures and/or when the matrix is less brittle. In this case extensive volume damage is done before the crack propagates and its extension by void coalescence involves further ductile matrix deformation (microshear banding). This mechanism leads to a rather rough zone of subcritical crack growth (Fig. 5(c)). The energy consumption and thus the toughness of the material *decrease* from Fig. 5(a) to Fig. 5(c) (Smith 1989, Smith *et al.* 1988).

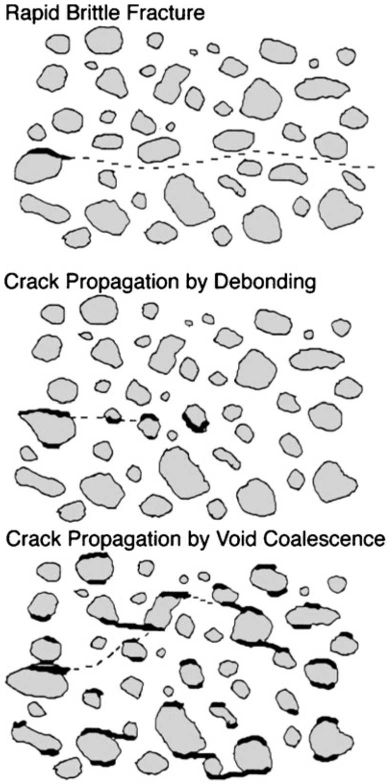

Rapid Brittle Fracture
Crack Propagation by Debonding
Crack Propagation by Void Coalescence

Figure 5
Schematic representation of crack propagation mechanisms in a highly filled epoxy resin. See text for details (after Smith 1989).

In compressive loading (of the same quartz flour-filled epoxy resin) matrix shear, particle–matrix debonding, particle fracture, and matrix microcracking by linking up of debonded zones are observed in this order. The latter three mechanisms are irreversible, and create the volume damage that is first observed close to the yield point.

The presence of well-bonded particles influences the toughness, K_{Ic}, positively (by crack pinning and bowing, void formation, and crack bridging). For many systems K_{Ic} and E are seen to increase linearly with Φ_f up to moderate filling fractions (Spanoudakis and Young 1984, Moloney *et al.* 1987). The linear relationship between K_{Ic} and E points to a crack opening displacement (COD) criterion for crack propagation (Spanoudakis and Young 1984). In comparing different filler materials one observes that K_{Ic}

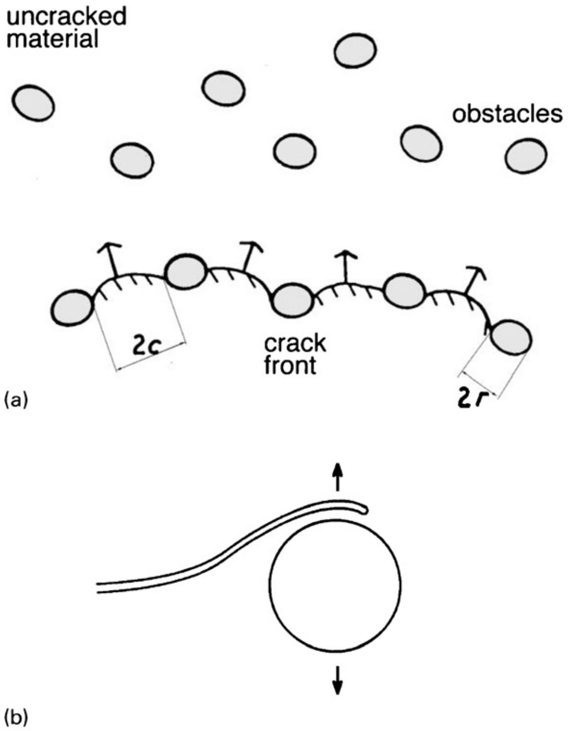


Figure 6
Modeling of the reinforcing effect of well-bonded particulate charges on subcritical crack propagation in filled thermosets through (a) crack pinning (Lange 1970, Evans 1972) and (b) crack deviation (Spanoudakis and Young 1984).

depends on the specific surface area of the fillers and very little on interfacial adhesion (Spanoudakis and Young 1984, Moloney *et al.* 1987, Nakamura and Yamaguchi 1993). In the case of rather weak filler materials (such as dolomite) toughening will be limited by transverse particle fracture. The energy release rate, G_{Ic}, of a filled resin generally is also larger than that of the pure resin; however, the relationship between G_{Ic} and Φ_{f} depends on filler size and interfacial adhesion. A strong interfacial bond is seen to be not always beneficial for a large G_{Ic}.

The notable increase in toughness of particulate-reinforced resins has been associated with the crack pinning mechanism (Spanoudakis and Young 1984, Moloney *et al.* 1987, Cantwell *et al.* 1988, Cantwell and Roulin-Moloney 1989, Lange 1970, Evans 1972). Well-bonded, rigid particles act as an obstacle and force the crack to bow (or to deviate) as it passes between neighboring filler particles (Fig. 6).

Evans (1972) has extended the line tension concept used in dislocation theory as suggested by Lange (1970). He has related the increase in toughness to the ratio of particle size, *r*, to spacing, *c*. Silica and alumina composites correspond quite well to the approximation thus obtained for rectangular particles (Moloney *et al.* 1987, Cantwell and Roulin-Moloney 1989, Evans 1972).

5. Conclusion

The formation of composites by blending filler particles of low aspect ratio with thermoplastic and thermosetting polymers leads to several beneficial effects. Young's modulus is considerably increased through active fillers of large specific surface area and through their influence on the structure of the surrounding matrix material (crystallinity, density). Time-dependent moduli can be derived from static ones by the separation of variables. The creep resistance is significantly increased, and the time–temperature–stress reduction principle has been applied successfully. To produce composites with adequate strength, fillers must show good bonding, but even then strength generally decreases with filler content. Toughness is less sensitive to the quality of adhesion, such that particle–matrix debonding (causing void formation) can be tolerated provided the formation of voids is well distributed in space and time. Apart for cost reduction, fillers are advantageously used if high strength, dimensional stability, and/or special thermal, optical, and electrical performances are required. Automotive and tribological components, lightweight garden furniture, opalescent packaging films, and high-voltage circuit breakers are some examples.

Bibliography

Bartczak Z, Argon A S, Cohen R E, Weinberg M 1999 *Polymer* **40**, 2347

Beer H R, Kaiser T, Moloney A C, Kausch H H 1986 *J. Mater. Sci.* **21**, 3661

Cantwell W J, Moloney A C, Kaiser T 1988 *J. Mater. Sci.* **23**, 1615

Cantwell W J, Roulin-Moloney A C 1989 In: Roulin-Moloney A C (ed.) *Fractography and Failure Mechanisms of Polymers and Composites*. Elsevier, London

Eurofillers 1999 *Eurofillers Int. Symp.*, Lyon

Evans A G 1972 *Phil. Mag.* **26**, 1327

Fekete E, Molnar Sz, Kim G-M, Michler G H, Pukanszky B 1999 *J. Macromol. Sci.: Phys.* **B38**, 885

Friedrich K 1978 *Progr. Colloid Polym. Sci.* **64**, 103

Galeski A 1995 In: Karger Kocsis J (ed.) *Polypropylene, Structure, Blends and Composites*. Chapman and Hall, London, Vol. 1, p. 1

Gregory A, Vogel D, Béguelin Ph, Gensler R, Kausch H H, Jaeger M, van Buren A 1999 *Mechan. Time-Dependent Mater.* **3**, 71

Jancar J, DiAnselmo A, DiBenedetto A T 1992 *Polym. Eng. Sci.* **32**, 1394

Joppien G R 1978 *Angew. Makromol. Chemie.* **70**, 189

Kaiser T, Cantwell W J, Smith J W, Moloney A C 1989 *Swiss Mater.* **1**, 17

Karger Kocsis J 1995 In: Karger Kocsis J (ed.) *Polypropylene, Structure, Blends and Composites*. Chapman and Hall, London, Vol. 1, p. 142

Katz H, Milewski J V (eds.) 1987 *Handbook of Fillers and Reinforcements for Plastics*, 2nd edn. Van Nostrand Reinhold, New York

Kausch H H, Gensler R, Grein Ch, Plummer C J G, Scaramuzzino P 1999 *J. Macromol. Sci.: Phys.* **B38**, 803

Kim G-M, Michler G H 1998a *Polymer* **39**, 5689

Kim G-M, Michler G H 1998b *Polymer* **39**, 5699

Kinloch A J, Young R J 1983 *Fracture Behaviour of Polymers*. Applied Science, London

Labour Th 1999 Thesis no. 99 ISAL 0012, Institut National des Sciences Appliquées de Lyon

Lange F F 1970 *Phil. Mag.* **22**, 983

Liang J Z, Ly R K Y 2000 *Polym. Int.* **49**, 1770

Michler G H 1992 *Kunststoff-Mikromechanik*. Hanser, Munich, Germany

Mineral Fillers in Thermoplastics 1999 Advances in Polymer Science. Springer, Berlin, Vol. 139

Moloney A C, Kausch H H, Kaiser T, Beer H R 1987 *J. Mater. Sci.* **22**, 381

Nairn J 1997 Invited lecture. *Eurofillers Int. Symp.*, Lyon; see also *Trans. ASME* 64, 804

Nakamura Y, Yamaguchi M 1993 *Polym. Eng. Sci.* **33**, 279

Nicolais L, Narkis M 1971 *Polym. Eng. Sci.* **11**, 194

Nielsen L E, Landel R F 1994 *Mechanical Properties of Polymers and Composites*. Dekker, New York

Petermann J 1995 In: Karger Kocsis J (ed.) *Polypropylene, Structure, Blends and Composites*. Chapman and Hall, London, Vol. 1, p. 1

Pukanszky B 1990 *Composites* **21**, 255

Pukanszky B 1993 *Makromol. Chem., Macromol. Symp.* **70/71**, 213

Pukanszky B 1995 In: Karger Kocsis J (ed.) *Polypropylene, Structure, Blends and Composites*. Chapman and Hall, London, p. 1

Pukanszky B, Belina K, Rockenbauer A, Maurer F H J 1994 *Composites* **25**, 205

Rothon R (ed.) 1995 *Particulate-Filled Polymer Composites*. Longman Harlow, UK

Schlumpf H P 1993 In: Gächter R, Müller H (eds.) *Plastics Additives*. Hanser, Munich, Germany, Chap. 9

Schlumpf H P 1994 *Omya Kontakt* (172)

Sjoerdsma S D 1989 *Polym. Commun.* **30**, 106

Smith J W 1989 Ph.D. thesis, no. 792, Ecole Polytechnique Fédérale de Lausanne

Smith J W, Kaiser T, Moloney A C 1988 *J. Mater. Sci.* **23**, 3833

Spanoudakis J, Young R J 1984 *J. Mater. Sci.* **19**, 473, 487

Stricker F, Bruch M, Mühlhaupt R 1997 *Polymer* **21**, 5347

Tjong S C, Li R K Y, Cheung T 1997 *Polym. Eng. Sci.* **37**, 166

Van Mier J G M, Schlangen E, Verwuur A 1993 In: Huet C (ed.) *Micromechanics of Concrete and Cementitious Composites*. Presses Polytechniques et Universitaires Romandes Lausanne, p. 159

Vogel D, Cantwell W J, Scaramuzzino P, Kausch H H, Youd S F 1994 In: *Proc. ECCM-CTS 2*. Hamburg, p. 409

Vollenberg P H T, Heikens D 1989 *Polymer* **30**, 1656

Vörös G, Fekete E, Pukanszky B 1997 *J. Adhesion* **64**, 229

Walter R, Friedrich K, Privalko V, Savadori A 1997 *J. Adhesion* **64**, 87

Yan S, Yang D, Petermann J 1998 *Polymer* **39**, 4569

H. H. Kausch and Ph. Béguelin
EPFL, Lausanne, Switzerland

Polymer Matrix Composites, Recycling of

The debate on sustainable development and on the use of nonrenewable resources has wide-ranging implications for polymer-based materials. The polymer industry is often targeted for its perceived role in the "throw-away society," in part because of the high visibility of polymers in municipal solid waste. As a consequence, the environmental performance of polymers has become an important aspect of product development.

Recycling of polymer composites is a particular challenge since these are materials that have been developed to provide highly tailored specific properties (see *Composite Materials: Overview*# and *Polymer Matrix Composites: Matrices and Processing**). Recycling is one of several options for their reuse. To determine the most efficient revaluation strategy, the environmental impact of the material over its whole life cycle should be evaluated, taking into account all the stages of its production and use.

Figure 1 shows the evolution of the durability of a composite over its life cycle. Several key points determine the durability of polymer composites (see *Composite Materials: Environmental Effects**). First, their organic matrices tend to degrade under thermomechanical loading, and this effect is aggravated in the presence of oxidative agents, solvents, and radiation. They are also macromolecular assemblies that display time- and temperature-dependent properties. Finally, they are seldom in thermodynamic equilibrium, which results in aging phenomena and microstructural evolution during service life.

Over the life cycle of a composite degradation occurs at both high temperatures, during processing, and at low temperatures, during service. The cumulative influence of both regimes can be significant in closed-loop recycling.

The two dominant causes of degradation in polymer composites during processing are thermomechanical degradation of the matrix and fiber attrition. All commercial polymer matrices contain antioxidant additives, to protect against thermally activated oxidative degradation, which results in a decrease of molecular weight and/or cross-linking. Viscous or adiabatic heating of the polymer melt, due to the combination of intense shear rates and low thermal conductivity, further contribute to such degradation phenomena. In natural fiber composites, chemical degradation during processing can involve thermal degradation and solvolysis reactions such as hydrolysis, as reviewed by Bledzki *et al.* (1998) for thermoplastics reinforced with wood fillers.

In processes that submit the matrix to strong shear or elongational deformation, mechanical degradation of the matrix may occur in the form of unfolding and scission of the polymer chains. Also, in thermoplastic extrusion compounding and molding techniques,

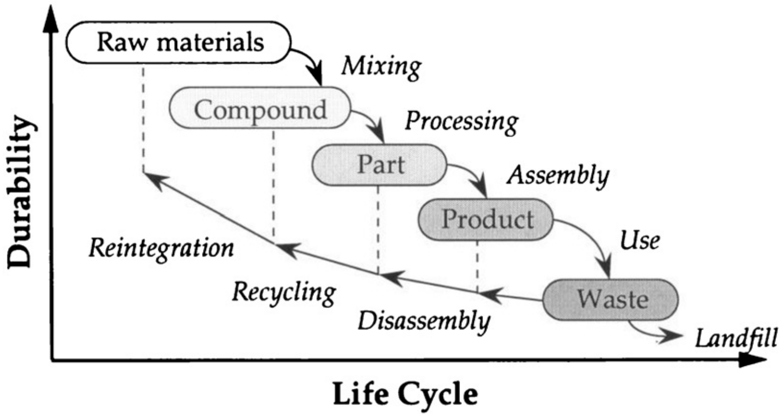


Figure 1
Polymer life cycle (Månson 1994).

fiber attrition is induced by the shear flow of the highly viscous polymer.

Thermo-oxidation is also a key feature of service-induced degradation of polymer composites. Tsotsis (1995) and Tsotsis and Lee (1998) reported comprehensive analyses of the thermo-oxidative degradation of carbon fiber-reinforced epoxy composites, which demonstrate the respective importance of matrix and interface degradation, matrix toughness, ply interaction, and edge effects. Furthermore, environmental aging combines the action of oxygen in the air, light, moisture, and temperature, and, possibly, mechanical loads. The action of oxygen appears to be the dominant factor in the evolution of the mechanical response of both reinforced-reinforced thermosetting and thermoplastic composites during long-term environmental exposure (Parvatareddy *et al.* 1995). Tant *et al.* (1995) reviewed the high-temperature properties and applications of polymers and polymer composites.

1. Recycling Technologies

The automotive and the electrical and electronics industries are the main sources of polymer composite waste streams, and the waste composites have interesting economic value for recycling. Although in Fig. 2 the principal recycling routes developed for plastics are shown, the techniques so far practicable for polymer composites are limited to the recovery of materials by chemical processes, the incorporation of regrind during compounding, and energy recovery through incineration.

1.1 Chemical Routes to Recycling

Chemical recycling is applied to the polymer matrix, and can be divided into four categories: solvolysis, oxidative processes, pyrolysis and similar thermal decomposition processes, and reductive processes. Solvolysis is a generic term for reactions with a solvent, and includes glycolysis, methanolysis, alcoholysis, ammonolysis, and hydrolysis. It is used mainly to depolymerize condensation-type polymers such as PET, PA, PMMA, and PC, and certain addition polymers such as polyurethanes, back into oligomers and original monomers, which can then be repolymerized into virgin materials. Addition polymers including polyolefins can be treated according to the three other processes. The potential and unique benefits of chemical recycling, and promising research directions, have been outlined by Tesoro and Wu (1995).

1.2 Mechanical Recycling

The mechanical recycling of polymer composites is evolving rapidly, driven by increasing legislation on waste treatment.

As a consequence of the reprocessing operation, process-induced fiber attrition and possible matrix degradation, as well as subsequent service-induced degradation and matrix aging after repeated process cycles, an inevitable reduction in material quality will occur. A thorough analysis of these effects is necessary to determine the durability and reliability of the recycled material.

Recycling of glass fiber-reinforced polyamide developed for under-the-bonnet automotive applications has particularly motivated research efforts. Dzeskiewicz *et al.* (1993) examined the recycling potential of glass reinforced-reinforced polyamides, and measured a systematic drop in mechanical properties (tensile strength, elongation, and impact) after multiple regrinding and injection molding operations. Eriksson *et al.* (1996a, 1996b, 1998) studied the

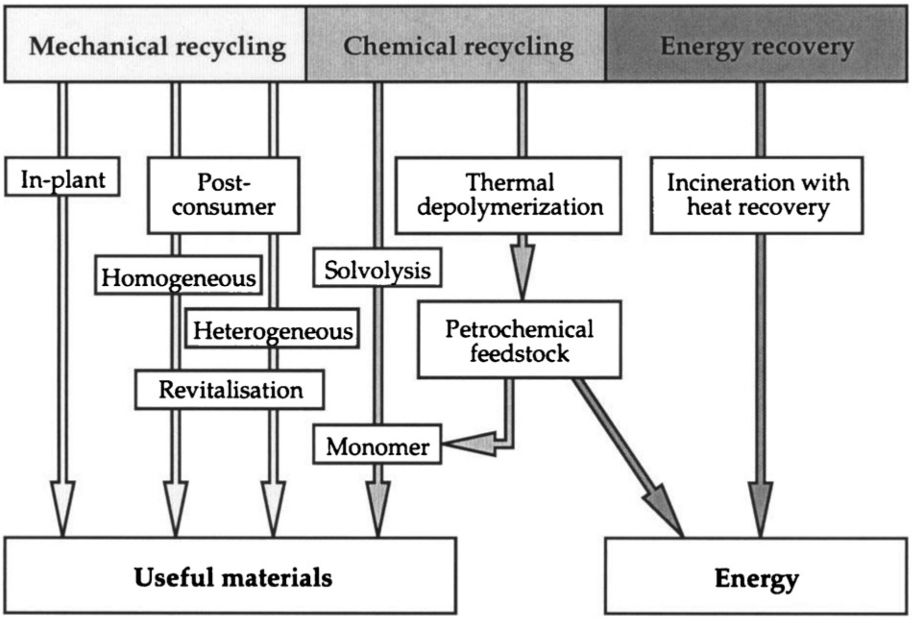


Figure 2
Recycling routes for polymer composites (after Lundquist *et al.* 2000).

recycling and durability of short glass fiber-reinforced polyamide 66. It was found that fiber length distribution, rather than matrix or interfacial thermal degradation, controls the short-term performance of the recycled composite. Fiber attrition is significant during initial compounding and first injection molding, but is less severe during further regrinding and remolding. The deterioration rate of recycled composites containing up to 25 wt.% regrind is similar to that of virgin samples during thermal aging but slightly faster during coolant aging. Critical sizes and concentrations of impurities are defined, below which a safe mechanical behavior is achieved.

Chu and Sullivan (1996) showed that recycled reinforced-reinforced polycarbonate possesses comparable or even better properties than the comparable commercial grade. Injection and extrusion compression molding yields recycled composites with good tensile properties, at the expense of impact strength, whereas the opposite is true for compression-molded samples, due to corresponding fiber orientation distributions.

As an alternative to mechanical grinding, which is systematically associated with mechanical recycling, selective dissolution of the polymer fraction has been proposed, followed by filtration under pressure to recover separately the fibers from the polymer solution (Papaspyrides and Poulakis 1997, Papaspyrides *et al.* 1995). The technique allows the removal of contaminants and degraded species such as cross-linked molecules.

To reuse thermoset resins and their composites, shredding and grinding the material into flakes or powder fractions with controlled sizes is necessary. Kelderman (1995) showed that these techniques are capable of generating a large variety of regrind fractions containing defined fiber length distributions. Whereas initially regrind was mostly used as filler, either in uncured thermoset resins or in virgin thermoplastics, its potential as a reinforcement has subsequently been favored. Major efforts have been directed to recycling sheet molding compounds (SMCs) and bulk molding compounds (BMCs), which account for about 50% of all composites used.

1.3 Incineration and Energy Recovery Routes

Incineration significantly reduces the volume of waste materials and, as such, has been promoted as an alternative to landfill, nowadays prohibitively expensive and subject to negative public opinion and legislation. Both of these waste management routes lead to a total loss of the material value, and are not free from environmental impacts. Incineration associated with energy recovery is, however, receiving increasing support as a viable recycling method.

Most of the efforts devoted to energy recovery of polymer composites have involved automotive shredder residues. Since 1990 the SMC Automotive Alliance, with 30 suppliers of the automotive industry, has promoted the combination of pyrolysis processes with mechanical recycling as the most realistic mid-term disposal options for ASR (e.g., *Automotive Engineering* 1994). Pyrolysis for recycling SMCs as an alternative to mechanical recycling was investigated by Demarco *et al.* (1997); it was found that temperatures in the range 400–500 °C are the most suitable, and for a glass fiber-reinforced orthophthalic polyester SMC, yields of 9–16 wt.% of nonpolluting liquid fuels with a high gross calorific value can be obtained, which can then be used alone or mixed with fuel oils.

2. Conclusion

The polymer industry is under strong pressure to provide viable recycling scenarios for their materials, and as a result life cycle issues have become an integral part of product design. The recycling of polymer composites, which are inherently tailored materials, requires tailored solutions: an optimum recycling strategy must be defined for each specific material combination.

In spite of many advances, much work is still necessary. A key aspect of future success will be the comprehensive evaluation not only of technical factors but also of economic, organizational, and legislative considerations.

Bibliography

Automotive Eng. 1994, **102**, 29–31

Bledzki A K, Reihmane S, Gassan J 1998 *Polym. Plast. Technol. Eng.* **37**, 451–68

Chu J, Sullivan J L 1996 *Polym. Compos.* **17**, 556–67

Demarco I, Legaretta J A, Laresgoiti M F, Torres A, Cambra J F, Chomon M J, Caballero B, Gondra K 1997 *J. Chem. Technol. Biotechnol.* **69**, 187–92

Dzeskiewicz L, Farrel R E, Winkler J 1993 *Eng. Plast.* **6**, 416–25

Eriksson P A, Albertsson A C, Boydell P, Eriksson K, Månson J A-E 1996a *Polym. Compos.* **17**, 823–9

Eriksson P A, Albertsson A C, Boydell P, Eriksson K, Prautzsch G, Månson J A-E 1996b *Polym. Compos.* **17**, 830–9

Eriksson P A, Albertsson A C, Boydell P, Månson J A-E 1998 *Polym. Eng. Sci.* **38**, 348–56

Kelderman H 1995 *Proc. R '95*. Geneva, Vol. I, pp. 324–31

Lundquist L, Leterrier Y, Sunderland P, Månson J A-E 2000 *Life Cycle Engineering of Plastics: Technology, Economy and the Environment*. Elsevier, Amsterdam

Månson J A-E 1994 New demands on manufacturing of composite materials. In: Chawla K K, Liaw P K, Fishman S G (eds.) *High Performance Composites—Commonalty of Phenomena*. TMS, Warrendale, PA, pp. 1–19

Papaspyrides C D, Poulakis J G 1997 *Proc. R '97*. Geneva, Vol. IV, pp. 120–5

Papaspyrides C D, Poulakis J G, Arvanitopoulos C D 1995 *Resources Conserv. Recycl.* **14**, 91–101

Parvatareddy H, Wang J Z, Dillard D A, Ward T C, Rogalski M E 1995 *Compos. Sci. Technol.* **53**, 339–409

Tant M R, McManus H L N, Rogers M E 1995 *ACS Symp. Ser.* **603**, 1–20

Tesoro G, Wu Y 1995 In: Rader C P, Baldwin S D, Cornell D D, Sadler G D, Stockel R F (eds.) *Plastics, Rubber and Paper Recycling. A Pragmatic Approach.* American Chemical Society, Washington, DC, pp. 502–10

Tsotsis T K 1995 *J. Compos. Mater.* **29**, 410–22

Tsotsis T K, Lee S M 1998 *Compos. Sci. Technol.* **58**, 355–68

P. Sunderland
Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland

---

Polymer Matrix Composites: Applications

The successful introduction of polymer matrix composite structures in competition with well-established and optimized solutions in metal, wood, or concrete, is a complex process. To varying degrees, depending on the economical sector, for a composite to replace an existing alternative material requires dual satisfaction of economical and technological requirements. Composite materials feature significant advantages over conventional materials in several respects. For example, their stiffness and strength are high while their density is low, their internal architecture and anisotropy can be controlled towards the implementation of innovative functional and structural design concepts, and their corrosion resistance is often excellent. They are also attractive materials from the economical standpoint: their fabrication processes allow for the production of complex shapes, raw materials can be inexpensive (e.g., when using reinforcements such as glass fiber), and their use can generate significant savings during service of the components produced.

Nonetheless, materials substitution generally proves a long and complex process; with composites this process can be rendered more so. Designing with composites is often a challenging task, because designing composite products means designing the materials too, and because composite properties are often highly anisotropic. The wide array of possible constituent materials and fiber architectures, and the wide range of specific applicable manufacturing techniques, further increase the relative difficulty of materials choices with composites. Economical advantage is also difficult to establish, and depends on the viewpoint adopted, in particular on whether the sale price of the component is the sole consideration or whether, on the contrary, the entire lifecycle of the component is analyzed. Finally, with composites, manufacturing is an important issue, in which the production of defect-free material at low cost still can be challenging.

Despite these challenges, composites are now part of everyday life, and have entered nearly all major industrial sectors, including aerospace, ground transport, packaging, the sports industry, and civil engineering. Most applications of composite materials are modern; however, some are in fact quite ancient.

1. Composites as Ancient Engineering Materials

Since prehistoric times the application of composite materials and the development of laminated structures have been related to local evolutions of craftsmanship. Wattle and daub or straw-reinforced clay became important building materials for permanent and more protective settlements. The earliest laminated composite structures, like lightweight composite bows and chariots, were glued layered structures of wood, hide, bone, sinew, and horn (Fig. 1).

The use and combination of those natural materials, polymeric and fibrous, based on chains of protein, cellulose, collagen, and keratin was a development made by prehistoric people who understood that by a proper combination of different materials they could explore their typical beneficial properties, while the poor capacities and drawbacks of the individual components often disappeared when these were combined into a composite material (Gordon 1991, Keegan 1993).

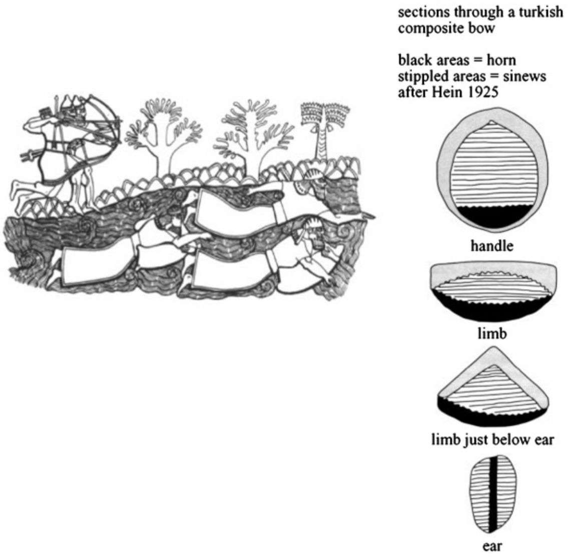


Figure 1
Ancient composites. Far Eastern composite bows may have been in use about 5000 years ago and reached perfection about 200 BC. Their design did not change until they became obsolete as a war weapon in the nineteenth century (Mesopotamian stele). The composite bow represents the first advanced (pretensioned) laminated composite machine. The bow was fabricated from a core of laminated wood steamed into a curved shape, to which were glued (glue made of boiled out fish skins, hides, and bones) on the outer side lengths of elastic animal (goat) tendon and on the inner side strips of compressible (ox) horn. Sensitivity towards delamination was eliminated by encapsulating the laminated structure with wet tendon, which shrinks after drying.

2. Modern Polymer Matrix Composites: Design Benefits and Considerations

With modern synthetic polymer matrix composites the high mechanical performance of fibers and their environmental resistance provide the basic ingredients towards the production of materials with outstanding properties for structural applications, particularly where low weight is desired. By combining such fibers with a polymeric matrix, the low mechanical properties of unreinforced polymers and the poor compressive strength of unbonded fibers essentially disappear. While in ancient times and also in military applications the implementation of composites has derived chiefly from performance improvement, generally the use of frequently expensive polymer composite materials can only be justified in designs where added value in function(s) and performance or by cheaper manufacturing leads to competitive price–performance ratios. Therefore, successful development and design of composite engineering structures must be guided by maximizing functional added value and minimizing the use of materials, energy, and cost over a lifetime.

Compared to ‘‘traditional’’ metal designs, modern well-designed and manufactured composite structures (as opposed to structures which duplicate ones designed for metals) offer higher mass-specific stiffness and strength, and better resistance to fatigue and environmental degradation (e.g., aircraft propeller blades, helicopter and windmill rotor blades). When crash or impact energy absorption is a leading design parameter, composite solutions can become vastly superior (e.g., Formula 1 racing cars, trainfronts, armor). When damage resistance or damage tolerance criteria are dominant concerns, even laminated polymer matrix composites become, if correctly designed and manufactured, superior to most steel and aluminum designs without weight penalties. Examples of this are provided by the wings and fuselages of modern military fighters and commercial aircraft.

Of all, however, the biggest potential benefits of polymeric composite materials are found in the possibility of improving the structural efficiency and effectivity (structural performance per unit weight and per unit cost) by almost unlimited geometrical and material ‘‘shapeability,’’ and the possibility of integrating the required mechanical, physical, and chemical properties of the structure using proper material design. Examples of this are relatively abundant, and include the B2 stealth bomber (Fig. 2), sailplanes, pressure vessels, temperature-conditioned transport, packaging, bridges, and canal locks. Awareness among future designers of functional and structural designs using composites, no matter for which product groups, of which fibers and polymers are applicable, including the possibilities offered by the many possible different "fiber architectures" and in relation to all available manufacturing techniques, are important factors in the engineering application of modern polymer matrix composites (Beukers and van Hinte 1998).

3. Polymer Matrix Composite Products: The Crucial Role of Industrial Manufacturing Techniques

In the nineteenth century the rapid introduction of electricity in communities and somewhat later of modern wheeled transport was made possible by the industrial capabilities of transforming natural fibers and fillers, phenolic polymers, and natural rubber into insulating structures for cables, switches, and electrical equipment. Automotive transport was revolutionized by the production of low-noise, lightweight, inflatable wheel tires made of fiber-reinforced natural rubber structures. Although labor was comparatively inexpensive, the successful breakthrough of these applications relied on the development of industrial techniques capable of ever-growing mass production. Ultimately, the parameter that has the most significant effect on the price–performance ratio of commercial products is the manufacturing cost; helmets or liquid petrol gas tanks provide examples of this. Once new materials and products have been accepted and established, the performance per unit weight gains importance and must be optimized. Decreasing structural "dead" weight is, of course, beneficial for performance improvement: this is applicable in packaging, transportation, and in "gravity"-limited structures such as tall bridges and buildings. In addition, minimizing the quantity of material leads to reductions in material cost, production time, and consequently in the cost of manufacturing (a liquid petrol gas tank is a good example (Fig. 3)). Increasingly, the success of polymer matrix composites in applications, by volume and by number, can be ranked by accessibility and reproducibility of the applied manufacturing techniques (Miravete *et al.* 1999).

3.1 Short and Discontinuous Fiber-reinforced Polymers

The introductory (pioneering) period for processes like injection molding and flow and form pressing of short fiber-reinforced polymers was based on thermosetting polymers, ranging from phenolics, polyesters,

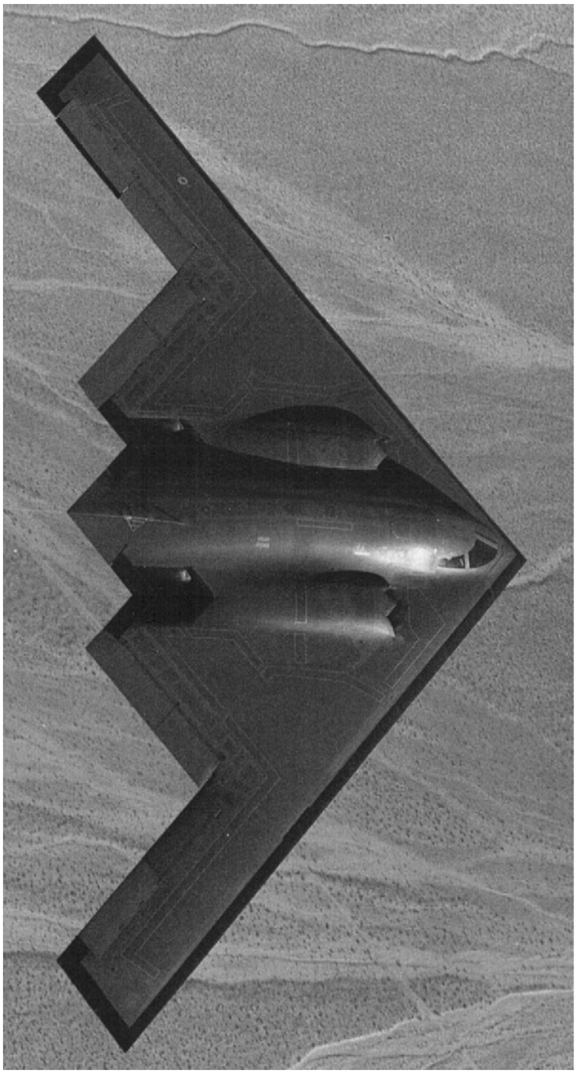


Figure 2
Polymer matrix composite in an aerospace application. The Northrop B-2 Spirit bomber is a new paradigm in aircraft design: 80% of its skin structure is made of carbon and glass fiber-reinforced thermosets. Its aerodynamic, functional, and structural design is trend setting for future aircraft design in general. Using the shaping possibilities and the mass-specific mechanical properties of carbon fiber-reinforced thermoset polymers has resulted in new structure development including more efficient flying wing configurations and improved stealth technology, and has given new impetus to aerodynamics (three-dimensional blended body aerodynamics) and stimuli for active stability and control solutions (tailless configurations). (Courtesy of Northrop Grumman.)

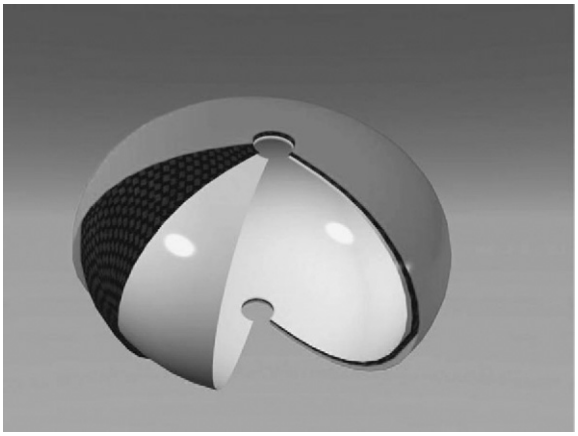


Figure 3
Polymer matrix composite in packaging. A wheel-shaped ultralight composite pressure vessel designed to store liquid petrol gas for cooking or car driving in urban areas. Pressure vessel design in polymer fiber composites can lead to more functional forms, like this carbon fiber-reinforced container (8550 g, 57 l). Separating the functions, like impermeability for gas and liquids, pressure tightness, and protection against impact and fire, makes possible a manufacturing strategy that can compete with metallic pressure vessel manufacturing. (Courtesy of Delft University/Advanced Lightweight Engineering.)

vinylesters, to epoxies. In the case of injection molding with short fibers (lengths less than about 10 mm, glass and sometimes carbon fibers) and in matched-die pressing with longer, discontinuous fibers (lengths less than about 100 mm, mostly glass fibers), thermosetting polymers are gradually being replaced by more expensive but technically equivalent or superior thermoplastics. The main reason for this is the cost reduction that results from cycle time reduction, which is crucial for mass-produced products. Fiber reinforcement is limited with respect to length, fiber content by volume (below 35%), and orientation control. Injection-molded products are generally small, but they can be complex in shape and produced with very close tolerances. A good example of the application of such composites is in camera housings.

Matched-die flow molding of larger plate and shell structures has also gained in importance. Stacked prefabricated sheets of discontinuous fiber-reinforced thermoset polymers (glass–polyester, sheet molding compound ("SMC")) or thermoplastic polymers (polypropylene, glassmat-reinforced thermoplastics ("GMT")) are formed and shaped by pressing at the right temperatures into medium-sized products. Compared to injection-molded composites, the length (shorter than 100 mm) and volume percentages (below 45%) of the reinforcing fibers, typically of chopped E-glass, are increased resulting in stronger reinforcement of the matrix and higher composite in-plane stiffness and strength. The control of fiber orientations is similar to injection molding and based more or less on keeping the fiber distribution as random and uniform as possible, in this case mainly in the in-plane directions. Typical applications of products of this processing route include parts for trucks and cars, as well as garden furniture.

For both techniques, heavy machinery is indispensable, rigid heated (cooled) molds are kept closed or are closed with a force sufficient to overcome the pressure of the process (for injection molding <1000 bar and for pressing 100–200 bar) multiplied with the projected part area, which can frequently be rather high. To cover the ensuing investment costs, large numbers of the composite parts have to be sold. Therefore, automotive components, housings for cameras, popular electrical and electronic equipment, furniture, and leisure and sporting goods have become popular applications for this class of polymeric matrix composites, typically displacing metal and wood.

3.2 Continuous Fiber-reinforced Polymer Matrix Composites

The short and discontinuous fiber-reinforced thermosets and thermoplastics are responsible for the biggest share of successful applications of polymer matrix composite materials, whether measured by number of parts or quantity of material used. Less visible, however, but growing enormously since the early 1990s, are the applications of continuous fiber-reinforced polymers. Instead of mass-produced polymer-based products, the applications of continuous fiber-reinforced composite parts are in general in advanced, capital-intensive, thermoset-based materials and products. Except for fast (reciprocating) machine parts and expensive and fashionable leisure and sporting goods (alpine sports, motorcar racing, angling, tennis, golf, and cycling), most of the applications of continuous fibers are bigger in size and take the form of plate and shell structures. For big structures (low numbers, large volumes) thermoset polymers such as phenolics, polyesters, and, increasingly, epoxies are mostly used, both for the relative low polymer cost per unit weight and the relatively low pressures and temperatures necessary for curing the structures.

In aerospace applications, continuous carbon fibers are increasingly replacing glass fibers to reduce structural weight, unless low material cost, good performance in impact protection, or specific electric or dielectric properties are required. For structural applications, up to temperatures of 200 °C, epoxies are mostly used as polymer matrix materials. In the case of interior applications, the domain of stringent fire, smoke, and toxic gas regulations, the phenolics are very popular. For low-weight, highly loaded structures a well-controlled and optimized fiber placement is essential. Designers and manufacturers often still depend on labor- or capital-intensive techniques with such composites. For example "laminating" can be done manually or by sophisticated automatic equipment, depending on the number of parts to be produced and on the design. In the stages of fiber (pre-)impregnation and in the cutting of patches of unidirectional tapes or woven fabrics, human labor is being replaced by cost-reducing and accurate machines. The most successful techniques in terms of composite material usage are filament winding of pressure vessels and pultrusion of composite profiles as intermediate structural elements applied in civil engineering: cables, bridges, and electricity masts provide examples of such applications. Except for carbon fibers in pretensioned cables, glass fibers are the most popular in this class of relatively large polymer composites. In pressure vessels, aramid and carbon fibers are becoming strong competitors, supported by the necessity of additional weight reduction and winding cycle time reduction to speed up production or to reduce the machine cost (Reinhart *et al.* 1987).

Also for continuous fiber-reinforced products, press forming is gaining in importance. For example, woven fabrics can in this process be draped into three-dimensional shape making use of a certain amount of thermoplastic binder material to keep the fabric "frozen" in the final shape. The result is a high-quality, preformed fiber structure that can be impregnated and cured in a following stage. The impregnation can be done in closed molds by resin transfer molding (RTM). Low-pressure vacuum-assisted RTM can be executed in simple (female) molds, the counter half or closing being just a flexible film material. In cases where additional pressure is used, relatively heavy and rigid matching dies are needed. Vacuum-assisted RTM is in particular becoming popular in the building of very large composite structures, such as hulls and decks for ships and windmill rotor blades, these structures having characteristic lengths up to about 30 m. "Pressurized" RTM is applied in the manufacturing of small propellers and helicopter rotor blades. Cost reduction and improvement of labor, product, and environmental quality are here combined when compared to traditionallaminating and impregnation techniques. For some applications, fabrics can be preimpregnated with a thermoplastic polymer (powder, melt, or solvent impregnation), and after heating these prepregs can be deformed (draped) and (re)consolidated in a mold. High fiber volume fractions and complete fiber orientation control is typical for this process and the speed is comparable to injection molding and other industrial forming processes used for short fiber-reinforced polymers. The pressure levels, however, are very low: for forming (draping) 1 bar or less is used, and for (re)consolidation 10–40 bar is needed. Helmets (Fig. 4) provide a good example of an application based on this process. When heating and cooling times are considered (a matter of seconds or minutes) it is clear that press forming of continuous fiber-reinforced composites is not only attractive because of the manufacturing speed, but also for the lightweight equipment and minimum energy required.

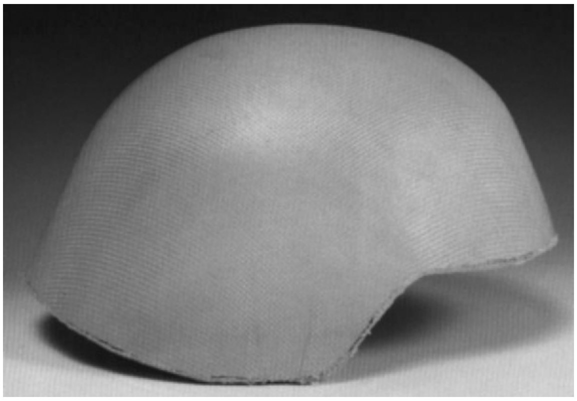


Figure 4
Polymer matrix composite in lightweight armor. Press formed (draped) aramid fiber-reinforced helmet, featuring improved functionality and weight reduction. Military helmets have traditionally been made out of steel, but like ancient (Mongol) multilayered textile armor, this helmet contains about 20 layers of plain woven aramid fibers, poorly bonded together (more than 65% fiber by volume) to absorb a maximum of ballistic impact energy. The latest thermoplastic polymer "blue" helmets manufactured by press forming (draping) are a compromise between protection against flying shrapnel and wearing comfort. They must be as light as possible (200 g less than their metal predecessors) and function in extreme cold and hot weather conditions alike. (Courtesy of Delft University/ Sumitomo.)

4. Conclusion

The scale of application of polymer matrix composites is very extensive. These structural materials have entered most major technological sectors including aerospace, automotive, leisure, consumer goods, electrical, electronic, and industrial manufacturing equipment, as well as many large-scale structures including ships and civil engineering structures. Factors determining the implementation of polymer matrix composites in engineering applications include both the intrinsic performance of the materials and the economic advantages that can be gained through their use. These factors, in turn, are strongly dependent on the somewhat specific nature of designing with composites, and also on the increasingly important role played by materials processes and part manufacturing procedures. Given the sustained rate of progress in these areas, it can be anticipated that composite materials will continue growing in importance in engineering for many years to come.

Bibliography

Beukers A, van Hinte E 1998 *Lightness, The Inevitable Renaissance of Minimum Energy Structures*. 010 Publishers, Rotterdam, The Netherlands

Gordon J E 1991 *Structures or Why Things Don't Fall Down*. Penguin, London

Keegan J 1993 *A History of Warfare*. Pimlico, London

Miravete, Bergsma D K, van Keulen F, Beukers A, de Boer H, Polynkine A A 1999 *3-D Textile Reinforcements in Composite Materials*. Woodhead Publishing, Cambridge, UK

Reinhart T J, Smith W S, Weidmer J C, Beckwith S C, Leslie J C, Quinlivan J T 1987 *Composites Engineered Materials Handbook*. ASM International, Materials Park, OH, Vol. I

A. Beukers

Delft University, The Netherlands

Polymer Matrix Composites: Matrices and Processing

Polymer matrix composites, and fiber-reinforced plastics (FRPs) in particular, are commonplace in the industrialized world; their applications range from primary structural aircraft components to tennis racquets, and many household appliances incorporate some form of fiber reinforcement. The development of engineering FRPs began once significant quantities of glass fiber became available in the 1940s, and was initially performance driven (e.g., for aerospace). Since then, reduction of materials costs has greatly widened their range of applications, and the impetus for continued development stems mainly from economic factors such as efficient fabrication and transformation, and a need for the energy savings possible with low-weight, high-strength, and high-modulus materials. However, FRPs have the major disadvantage with respect to metal- and ceramic-based structural materials that their service temperatures are limited to below 400 °C, even for the most advanced matrices.

An overview is given in this article of the most widely used FRP matrix materials and their relative merits. The various stages of processing to form composite parts are discussed, with reference to cost forecasting, life cycle engineering, and developing technologies. Finally, the important question of process–structure–property relationships in thermoset- and thermoplastic-based FRPs is addressed.

1. Matrix Materials

Apart from injection molding thermoplastics, which typically contain short (<1 mm) fibers, most FRPs are based on thermoset resins. Thermosets can be cured catalytically, thermolytically, or photolytically over a wide range of temperatures to give highly cross-linked solids. One of their principal advantages for FRPs is their low initial viscosity, which facilitates fiber impregnation, particularly where the base "A-stage" resin consists of monomeric or oligomeric components. Partial cure may then be used to produce solid "B-stage" prepregs used as precursors for a variety of processing routes. There is nevertheless great interest in thermoplastic polymers; although the high molar mass of thermoplastics implies that higher processing temperatures or specialized techniques are needed for impregnation, this may be outweighed by other factors: (i) they are generally nonreactive, implying an indefinite shelf-life for precursors; (ii) forming processes are reversible, facilitating recycling and repair; and (iii) cycle times are not limited by reaction kinetics, and mixing steps, which may involve toxic precursors and need careful control to ensure reproducibility, are eliminated. Table 1 summarizes the main trends, although it should be stressed that there is considerable overlap, the toughness of FRPs based on both classes of resins being highly sensitive to, for example, fiber morphology, interfacial adhesion, and the presence of additives (see *Composites: Interfaces**, *Continuous Parallel Fiber Composites: Fracture**). Moreover, each class of resins encompasses a wide variety of materials and materials properties, which will now be presented in more detail along with typical applications. A more systematic discussion of the applications of polymer matrix composites is to be found in *Polymer Matrix Composites: Applications**.

1.1 Thermoset Resins

There are several main types of thermosets, distinguished by their chemistry, curing mechanisms, and

Table 1
Property and process characteristics of thermosets and thermoplastics.

Property	Thermoset	Thermoplastic
Modulus	High	Medium
Service temperature	High	Medium
Toughness	Medium	High
Viscosity	Low	High
Processing temperature	Low	High
Recyclability	Limited	Good

final properties (Charrier 1990). Those discussed here are the most important for FRPs, namely polyesters, epoxies, and high-performance thermosets, although other thermosets, including polyurethanes (PUs), silicone rubbers, and phenol–formaldehydes (PFs), are also sometimes used as FRP matrices (e.g., PF–aramid composites for aircraft interiors).

(*a*) *Unsaturated polyesters*. Unsaturated polyesters (UPs) are the most widely used thermosets for FRPs. They consist of prepolymers cross-linkable by addition, initiated photolytically or by thermal decomposition of peroxides. Processing involves most of the methods described in the next section, and indeed UPs were originally developed for use in FRPs. However, to ensure adequate flow behavior, up to 50 wt.% styrene co-reagent is often added to give a so-called UPE resin. Environmental concerns have since led to the use of low styrene emission additives in UPEs. These segregate to, and seal, the surface during the final stages of cure, compromising adhesion in some cases. Another potential difficulty is excessive shrinkage, but this is largely avoided by the use of thermoplastic low profile additives. Once cured, unsaturated polyesters have heat deflection temperatures (HDTs) ranging from 80 °C to 150 °C in certain vinylester resins (the HDT is a softening temperature measured according to industrial norms and marks the approximate upper limit to the range of service temperatures). Their main composite applications include the construction industry, the transport industry (boats, car body panels), furniture, and chemical storage tanks.

(*b*) *Epoxy resins*. Epoxy resins may also be used in conjunction with most composite processing techniques, and are particularly suited to B-stage processing. Cure generally involves polyaddition with a multifunctional co-reagent, referred to as the "hardener," to produce a three-dimensional network with low shrinkage, and excellent adhesion and environmental resistance. Drawbacks of conventional resins include somewhat higher viscosities than in UPEs, long cure times, and moisture uptake. A wide range of resins and co-reagents is nevertheless available, products ranging from tough low-temperature epoxies to more brittle, high-temperature epoxies for advanced FRPs (prepregs for aircraft components account for around 30% of total production). Commonly used aliphatic amine hardeners permit room temperature cure, but the reaction is strongly exothermic, causing problems in thick moldings. Aromatic amine hardeners require higher curing temperatures but are better suited to large parts and give relatively high HDTs (up to 230 °C). Anhydride hardeners are less toxic and less polar than amines, but need an accelerator. With suitable catalysis, epoxies may also be cured thermally or photolytically without a co-reagent.

(*c*) *High-performance thermosets*. High-performance thermosets include allyls, cyanate esters, and liquid crystalline thermosets. Among the most successful are addition polyimides, developed mainly for the aerospace and electronics industries. Their HDT may reach 400 °C after cure, and they have excellent stiffness and environmental resistance. The cure mechanism is thermally induced free radical addition of a polyimide precursor via reactive end groups in the presence of peroxide, avoiding problems with the volatile by-products of conventional condensation reactions. Applications include jet and automobile engine components, heat-resistant panels, etc. However, some potential future applications of high-performance FRPs, such as reusable launch vehicles, will require HDTs in excess of 500 °C, setting the standard for the next generation of thermoset matrix materials.

1.2 Thermoplastic Resins

Both semicrystalline and amorphous glassy polymers form viscous melts that need to be processed at temperatures well above the glass transition temperature, T_g. However, since semicrystalline polymers melt well above T_g, it is the melting point, T_m, that defines the processing window. Whereas thermoset designations tend to be generic, most thermoplastics have well-defined linear architectures that undergo little change during processing. Their property profiles are generally adjusted for different processing routes by varying the molar mass or additive package (Charrier 1990); however, it should be borne in mind that other important properties such as toughness and ultimate tensile strength are also highly sensitive to changes in molar mass.

(*a*) *Semicrystalline thermoplastics*. Semicrystalline thermoplastics are tough and resistant to organic solvents, but many inexpensive and easy-to-process low T_g polymers, such as polyethylene, also show poor adhesion owing to their lack of functionality, and low softening temperatures, often associated with excessive creep. A good compromise between cost and performance is provided by polypropylene (PP, $T_m \approx 170\,^\circ\mathrm{C}$), for which functionalized additives are available to improve adhesion. PP is widely used in FRP applications in the automotive industry, for example, owing notably to glass mat thermoplastic technology. Polyamides (PAs or nylons), such as PA6 and PA6,6, are also popular, having higher HDTs than most thermosets, although they are moisture sensitive and more than twice as expensive as PP. Another polyamide of interest for FRPs is PA12, the versatile synthesis and functionality of which make it especially attractive for reactive FRP processing.

High T_g semicrystalline polymers such as poly(phenylene sulfide) (PPS) and poly(ether ether ketone) (PEEK) require correspondingly high processing temperatures, PEEK melting at about 345 °C, for example. However, these polymers have outstanding properties, including exceptional toughness and fatigue resistance. Particular effort has gone into PEEK–carbon fiber composite technology, leading to the so-called aromatic polymer composite (APC) prepregs, considered to be a major step in the development of advanced thermoplastic matrix FRPs (Cogswell 1992). However, PEEK is nearly two orders of magnitude more expensive than PP (Charrier 1990).

(*b*) *Amorphous glassy thermoplastics.* Aromatic high-performance amorphous thermoplastics such as poly(ether sulfone) (PES, $T_g \approx 230\,°C$) or poly(ether imide) (PEI, $T_g \approx 215\,°C$) have also attracted interest as advanced FRP matrix materials. They are less expensive than PEEK and tougher than thermosets, but there remain (arguably unjustified) concerns over their solvent resistance.

2. Processing

Figure 1 summarizes the various routes by which raw materials (resin plus fiber) are transformed into the final part (Månson 1994). A description of the most important of these now follows. Further details concerning specific types of reinforcement, and glass fibers in particular, are to be found in *Fibrous Reinforcements for Composites: Overview** and *Glass Fibers**.

2.1 Fibers, Prepregs, Feedstocks

Fibers are usually supplied in the form of rovings (glass fibers) or tows (carbon fibers). Rovings consist of straight continuous glass-fiber strands or bundles of about 200 filaments; the number of strands depends on the end use, and these may be several kilometers long. Tows are likewise available in various configurations. The fibers are typically sized (coated) during production to promote wetting and adhesion, silane coupling agents being widely used for glass fibers. They may be woven into fabrics if required, including *ad hoc* fiber preforms for specific applications. Chopped fiber lengths range from less than 1 mm for injection, to around 50 mm in randomly oriented mats for laminates.

Of major concern in FRP processing are trapped air and volatiles, since voids substantially reduce the final strength. Low-viscosity resins, vacuum outgassing, and consolidation all help reduce porosity. Use of preconsolidated sheets or prepregs also facilitates certain processes (see Fig. 1). These are available as impregnated fabrics or continuous warp sheet typically containing 60–65 vol.% uniaxially oriented fibers. Impregnation with thermosets may be achieved by dipping or hot melt processing (rolling of the fibers together with a resin film). When the viscosity is high, flow distances must be kept to a minimum, and intimate mixing of fine polymer powders or fibers with the fiber reinforcement often precedes melt processing of thermoplastic composites. Solvent processing is one alternative; it may lead to problems

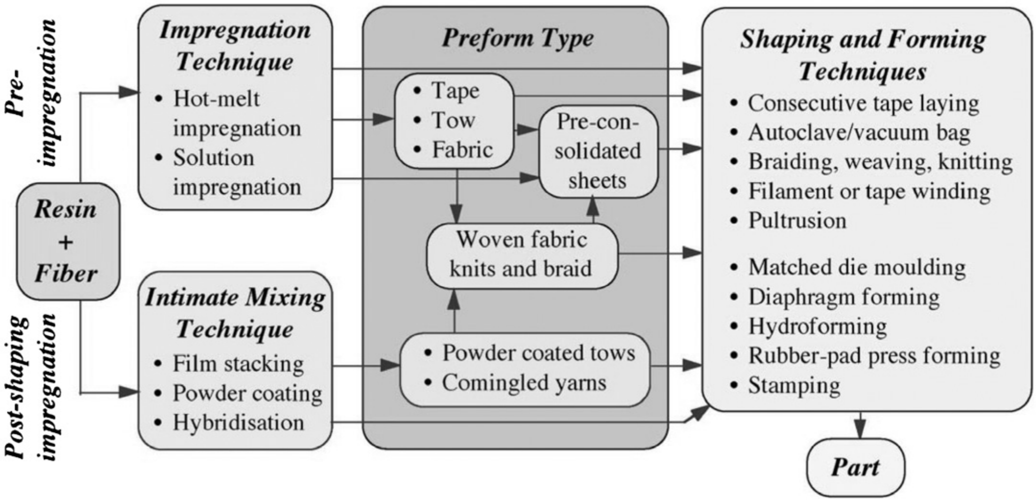

Pre-impregnation
Post-shaping impregnation
Resin + Fiber
Impregnation Technique
• Hot-melt impregnation
• Solution impregnation
Intimate Mixing Technique
• Film stacking
• Powder coating
• Hybridisation
Preform Type
• Tape
• Tow
• Fabric
Pre-consolidated sheets
Woven fabric knits and braid
• Powder coated tows
• Comingled yarns
Shaping and Forming Techniques
• Consecutive tape laying
• Autoclave/vacuum bag
• Braiding, weaving, knitting
• Filament or tape winding
• Pultrusion
• Matched die moulding
• Diaphragm forming
• Hydroforming
• Rubber-pad press forming
• Stamping
Part

Figure 1
Alternative manufacturing routes: from resin and fiber to finished part (after Månson 1994).

with residual solvent in the later stages of fabrication but has been used with success for PEI prepregs, for example. Another possibility is to chain extend relatively low molar mass precursors *in situ* by reaction of suitable end groups, an approach frequently used for high-performance thermoplastic FRPs.

Sheet and dough molding compounds are mostly based on unsaturated polyester resins containing filler and chopped fibers. Sheet molding compounds (SMCs) are produced from chopped strand mat placed between layers of filled resin containing thickening agents to promote moldability. Impregnation involves rolling of the resulting sandwich structure. SMCs usually contain 25–35 wt.% fibers but the fiber content can be up to 65 wt.%, in which case they are called HMCs. XMCs are similar to SMCs, but incorporate surface layers containing continuous fiber rovings. Dough molding compounds (DMCs), or bulk molding compounds (BMCs), contain shorter fibers than SMCs and are produced by mechanical mixing, which results in a more markedly three-dimensional fiber orientation distribution. Glass mat thermoplastics (GMTs) can be produced in a similar manner to SMCs, but the high matrix viscosity means that success is very dependent on the choice of fiber mat. Powder processing routes are an interesting alternative in this case.

It is sometimes advantageous to preassemble reinforcing fibers to fit a given shape, rather than rely on flow to distribute them throughout a mold. If displacement of the fibers during subsequent introduction of the resin is a problem, they can be anchored in place by spraying with a binder, or substituted by a preimpregnated fabric shaped into a suitable perform.

2.2 Shaping and Forming Techniques

The simplest forming technique is hand placement, where woven fabric or chopped strand mat is placed on a nonstick former and premixed liquid thermoset resin added with a brush, say, or by spraying. It is versatile and requires low investment, but the resulting mechanical properties are limited and the manufacturer may fall foul of legislation against volatiles. It is used for one-off or short production runs, including large objects, such as boat hulls. However, for reproducible, high-quality, cost-effective production other techniques must be considered.

(*a*) *Press molding.* Press molding using matched male and female tools is the most widely used technique for volume production of FRPs, being suitable for most types of precursor. With SMCs or DMCs, complex shapes can readily be formed at pressures of 3–15 MPa with cycle times of the order of minutes depending on the cure rate and the part size. It is also used for continuous fiber FRPs, but the limited flow restricts the practically attainable range of wall thickness and curvature. For mass production, steel tools and hence additional capital investment may be necessary, but for short runs, epoxy composite tools are often adequate. Variations include rubber-pad press molding, hydroforming, diaphragm forming, and cold stamping.

(*b*) *Vacuum and autoclave molding.* In vacuum molding, atmospheric pressure consolidates the material during cure, avoiding the need for a press. A prepreg or fabric laminate is placed on a mold, covered with a flexible membrane and the assembly sealed and evacuated. It provides a low-cost alternative to press molding for large, thin shells. If further consolidation is needed, cure may be done in an autoclave. Autoclaves usually operate at up to 1.5 MPa pressure and are widely used in the aerospace industry, having the advantage over press molding of cheap tooling for large parts and being ideal for complex structures such as honeycomb core sandwich composites. Indeed almost all prepreg material used in the aerospace industry is autoclaved. It is less suited to high-volume manufacture, although progress in automatic cutting and lay-up operations has led to considerable cost reductions and new market opportunities.

(*c*) *Resin transfer molding, reaction injection molding.* In resin transfer molding (RTM), the fiber reinforcement (either dry fibers or a preform) is placed in the mold cavity, and the resin then introduced by injection. RTM generally requires low-viscosity resins, but impregnation can also be improved by vacuum-assisted injection. It is used to produce large moldings such as car body shells, and the low injection pressures mean the molds themselves can be made from FRP, with considerable cost savings. However, although RTM is cheaper than prepreg processing, it is more limited in its range of application. Reaction injection molding (RIM) makes use of rapidly reacting components mixed just prior to injection, permitting very short cycle times, and short fiber reinforcement can also be introduced in one or both reagents (this is called reinforced reaction injection molding (RRIM)). It is competitive with press molding for the manufacture of a variety of automotive and domestic components.

(*d*) *Pultrusion.* Pultrusion is the continuous fiber FRP equivalent of extrusion, permitting continuous production of constant profile parts such as fishing rods, antennas, and channel profiles (or constant section parts, if the profile is modified *in situ*). Unlike in extrusion, however, the reinforcing fibers themselves are used to *pull* the material through a heated forming die, as shown schematically in Fig. 2. Rovings or tows are passed through a resin bath followed by a series of carding plates or ‘‘preformers,’’ which force

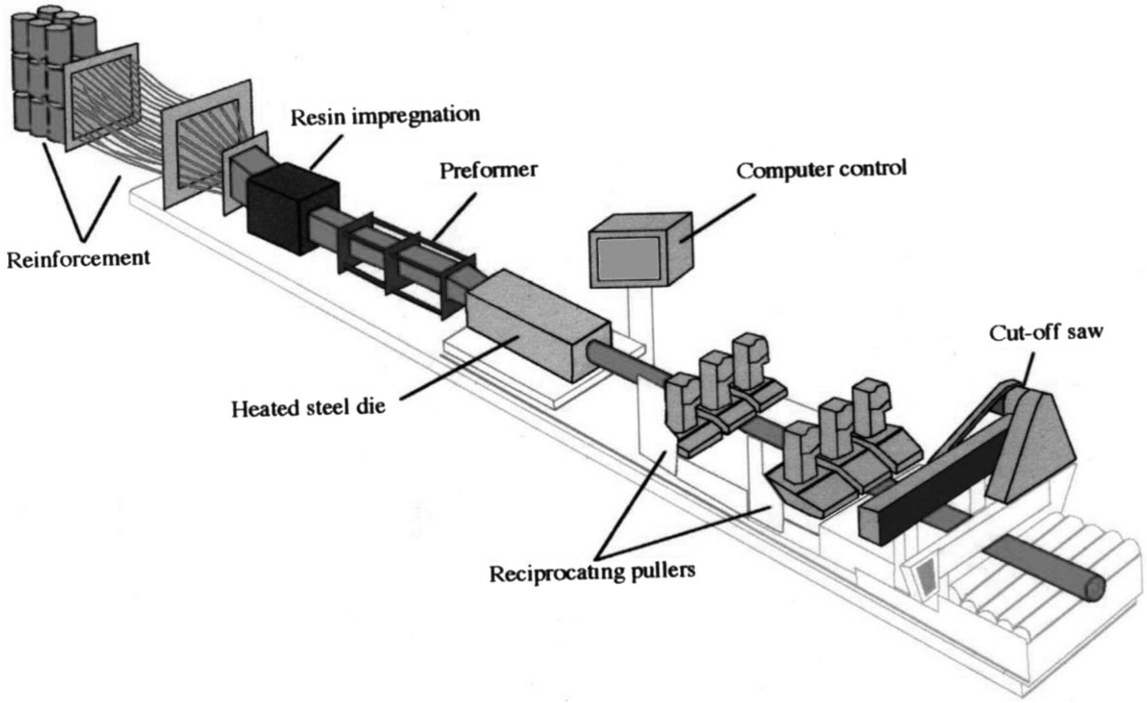


Figure 2
Schematic of the pultrusion process.

the resin into the fibers. The impregnated fibers are then pulled through the die, where the resin gels and cures. Preheating by radio-frequency methods may be used to increase throughput, resulting in production rates of up to several meters a minute. In some cases the resin is also injected directly into the die, eliminating the need for a preimpregnation step.

(*e*) *Filament winding*. Strong, stiff shell structures can be made by winding bands of continuous fibers impregnated with resin onto mandrels, which are removed after cure. Filament winding is widely used for the manufacture of reinforced plastic pipe and other cylindrical objects. However, modern filament winding techniques can also be used to produce complex structures such as aircraft fuselages. In such cases, the fibers must be wound along constantly changing nonrepetitive geodesic paths to prevent slippage. Automation of this technique has led to the development of advanced computer-based integrated design and manufacturing methods.

(*f*) *Developing technologies*. Some representative examples are now be given of developing technologies aimed at improving, or providing alternatives to established processing routes. The resin infusion process (RIFT) and its variants, in which a vacuum is used to draw resin into a preform or fiber bed (cf. RTM), has been successfully applied to large objects such as boat hulls. Large injection molding body technology (LIMBT), offers a viable alternative to press molding of SMCs and GMTs, contrary to the widely held notion that injection molding facilities become prohibitively expensive as part sizes increase; for example, it has been used to make large automotive body components from various thermoplastic FRPs (Argeropoulos *et al.* 1999). Integrated processing (IP) is aimed at improving manufacturing flexibility and adding value to composite products by combining different processes and materials; selective placement of reinforcement and integration of functions are thus achieved in short cycle times (Månson *et al.* 1995, Bourban *et al.* 1998). Reactive thermoplastics processing, developed for PA12-based composites, uses monomeric precursors for impregnation followed by polymerization *in situ*, avoiding the problems associated with the use of melt or solution impregnation techniques for thermoplastics (Schmid and Eder 1997, Bourban 2000).

2.3 Choice of Materials and Processing Routes: Composite Processing Maps

Figure 3 shows a map of various composite processing methods arranged according to the combination

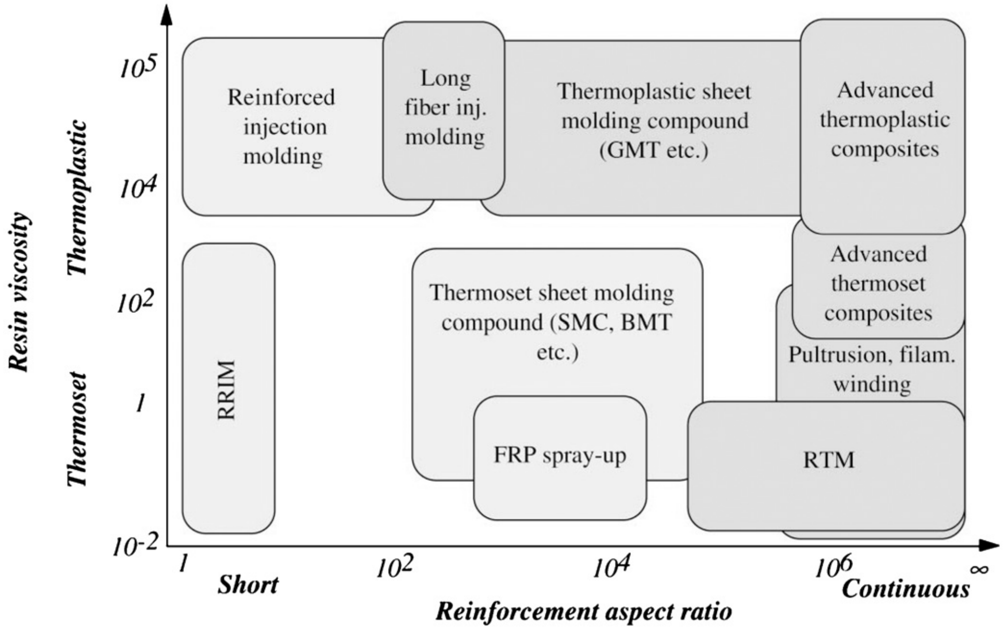

Reinforced injection molding
Long fiber inj. molding
Thermoplastic sheet molding compound (GMT etc.)
Advanced thermoplastic composites
RRIM
Thermoset sheet molding compound (SMC, BMT etc.)
Advanced thermoset composites
Pultrusion, filam. winding
FRP spray-up
RTM
10^5
10^4
10^2
1
10^-2
Resin viscosity
Thermoplastic
Thermoset
1
10^2
10^4
10^6
∞
Short
Continuous
Reinforcement aspect ratio

Figure 3
Composite processing map: viscosity and reinforcement aspect ratio (after Månson 1994).

of resin viscosity and reinforcement aspect ratio to which they are most suited. If viscosity is an overriding consideration it will generally dictate the choice between thermoplastics and thermosets, but it is immediately clear from Fig. 3 that the range of processes is broad enough for solutions to be available for thermoplastics or thermosets for any required reinforcement aspect ratio. Such "composite processing maps" may be based on other parameters, such as the potential for control of reinforcement orientation and aspect ratio; injection molding, for example, gives little orientation control and is limited to low aspect ratios, whereas three-dimensional control and long fiber orientation require techniques such as filament winding, hand lay-up and RTM/preform technologies. Figure 4 gives a further example that is particularly important for costing (see below), namely the relationship between the suitability of a given process route for complex shapes and its suitability for high-volume production, as defined by the cycle time.

2.4 Cost Modeling, Life Cycle Engineering

Since material costs are now substantially irreducible, and given that manufacturing often represents close to 60% of expenditure on FRP composite parts, the processing route is generally the most critical factor in cost-effective production. Assessing the economic viability of a processing route requires comparative cost analysis and hence reliable data and cost estimation methods. A variety of approaches for evaluating manufacturing costs have been proposed (Månson *et al.* 2000): comparative studies, process-oriented cost models, parametric cost models, and process flow simulation. Comparative studies are based on historical data and may suffice when envisaging the use of a standard technique for a component similar to those already in production. Process-oriented cost models are better adapted to new processes, and incorporate identification and quantification of cost drivers, but require detailed process information and a precise definition of the components. Parametric models are more versatile and allow easy manipulation of process and economic factors; however, they are based on the assumption of independent manufacturing steps, and may consequently underestimate costs. Such difficulties may be avoided by using process flow simulations, which take into account interactions between the different manufacturing steps, and provide forecasting tools that are also suitable for improvement studies, identifying bottlenecks and estimating repair costs. Given the growing concern for environmental issues, the successful introduction of a new product or process may also require establishment of a life cycle plan, particularly for the large-volume applications. As well as the impact of material on the environment during manufacture, service, and disposal, it is important to consider conservation of raw material value. This is a potentially important issue when choosing between thermosets and thermoplastics, since recycling of the former is severely limited. Furthermore, recovery from continuous fiber FRPs is difficult even for thermoplastic matrices. One alternative under consideration is recovery of the monomer via tertiary recycling (see *Polymer Matrix Composites, Recycling of**).

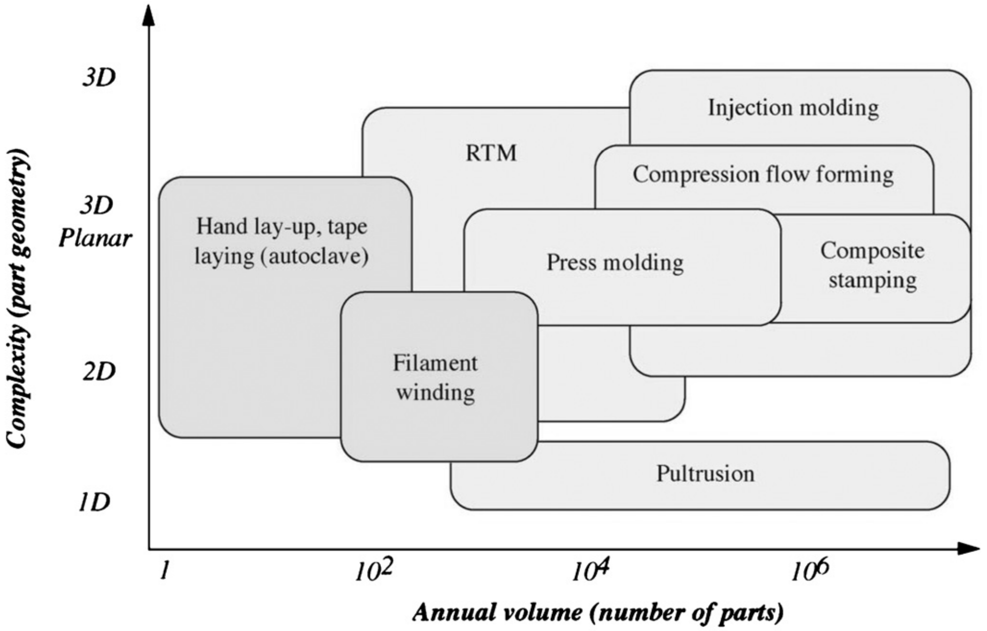

3D
3D Planar
2D
1D
Complexity (part geometry)
Injection molding
RTM
Compression flow forming
Hand lay-up, tape laying (autoclave)
Press molding
Composite stamping
Filament winding
Pultrusion
1
10^2
10^4
10^6
Annual volume (number of parts)

Figure 4
Composite processing map: complex shaping and volume production (after Månson 1994).

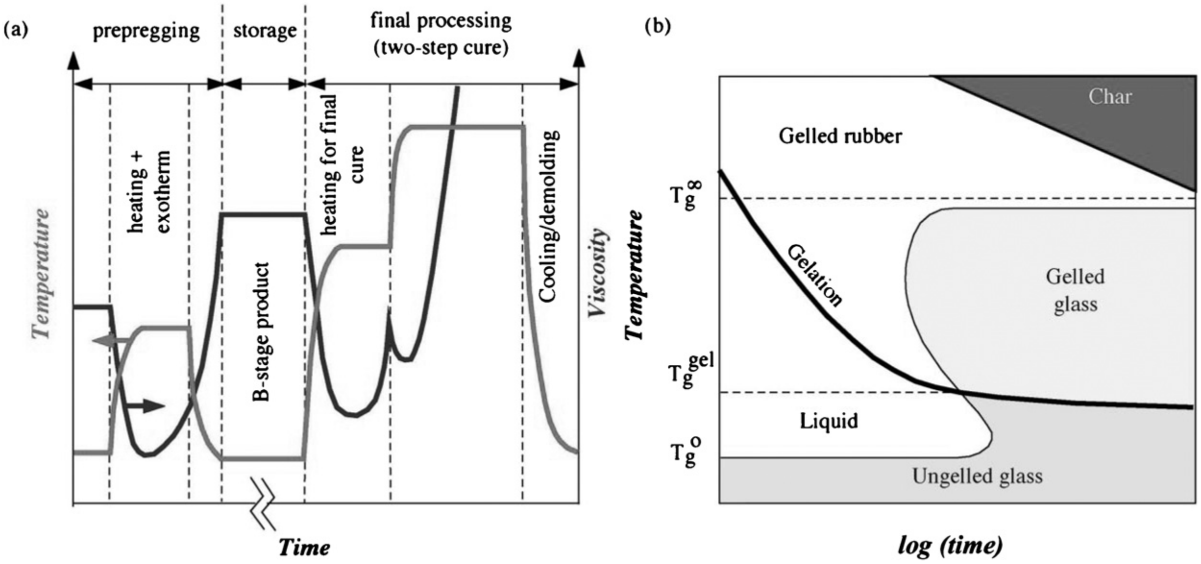


Figure 5
(a) Typical thermal cure history for a thermoset composite. (b) Isothermal TTT diagram associated with the cure of a thermoset polymer.

3. Process–Structure–Property Relationships

The thermal cure history shown in Fig. 5(a) is typical of prepreg and autoclave processing sequences for a thermoset FRP. During impregnation, the resin viscosity first decreases owing to the rise in temperature, then increases as cross-linking proceeds. The reaction is interrupted by cooling, and the resulting B-stage material is stored at low temperature to inhibit further reaction. During final processing, the viscosity again passes through a minimum, tending to infinity as the gel point is approached. A two-step final cure is often used to improve consolidation and dimensional stability, after which the part is cooled and removed from the mold or former. The final properties of thermoset FRPs are strongly dependent on the details of such time–temperature trajectories. Figure 5(b) shows a schematic time–temperature–transformation (TTT) diagram illustrating the main trends (Pang and Gillham 1990). During isothermal cure at temperatures between T_g^{gel} and T_g^{∞}, the polymer gels and then vitrifies, defining optimum conditions for final processing. Isothermal cure between T_g^{gel} and T_g^0, the initial glass transition temperature (or melting point), would cause the polymer to vitrify prior to gelation, hindering further reaction.

In a typical processing cycle for a thermoplastic FRP, the components are first heated to above T_m of the resin (or T_g for a fully amorphous resin). Pressure is applied for long enough to give adequate consolidation, and the resulting composite is cooled with the pressure maintained. Optimum processing conditions may be defined with the aid of processing windows, such as that shown schematically in Fig. 6. The lower bound for the cooling (and heating) rate is fixed by economic constraints on the cycle time. Slow cooling and heating may also result in prolonged exposure of the resin to high temperatures, leading to thermal degradation. At very high cooling rates, however, considerable thermal gradients may develop through the thickness of the component, leading to pronounced skin–core morphologies in semicrystalline resins, and excessive thermal stresses, as well as changes in primary morphological parameters such as spherulite size, lamellar thickness, and the degree of crystallinity. Increasing the hold pressure may help suppress defects resulting from internal stresses, but a maximum pressure will be imposed either by the performance of the press, or by the cost of high-pressure molds. High pressure may also damage the fiber bed and/or cause resin starvation.

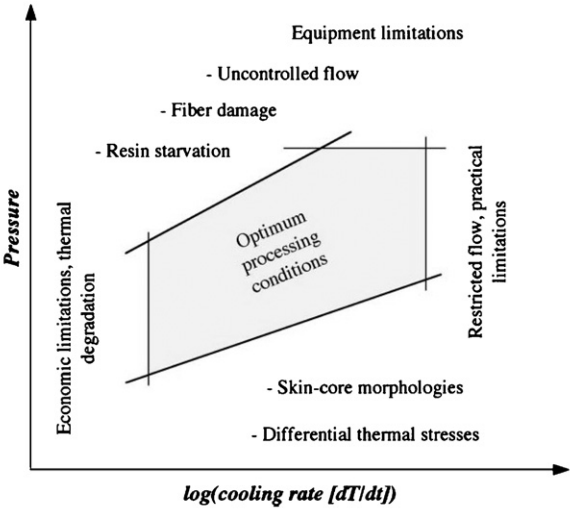

Equipment limitations
- Uncontrolled flow
- Fiber damage
- Resin starvation
Pressure
Economic limitations, thermal degradation
Optimum processing conditions
Restricted flow, practical limitations
- Skin-core morphologies
- Differential thermal stresses
log(cooling rate [dT/dt])

Figure 6
Generic processing window for a thermoplastic composite (after Månson 1994).

4. Conclusion

Although the initial driving force for the development of FRP composites was performance, economics is now the overriding consideration. Raw material costs are generally higher than for metals, but this is offset by weight reduction, increased design freedom, and reduced "part count" in a given application; even image can be a strong selling point for sports goods. Nowadays, it is innovation in processing that offers the greatest scope for increasing the engineering viability of polymer matrix composite materials, and for paving the way to new applications of these materials. Doubtless some of the techniques under development will shortly join or supplant the "standard methods" referred to here.

Bibliography

Argeropoulos J G, Fielding C R, Oswald L J 1999 *SAC 1999-01-3244*

Bourban P-E 2000 Liquid molding of thermoplastic composites. In: Kelly A, Zweben C (eds.) *Comprehensive Composite Materials*. Pergamon, Oxford, Vol. 2, pp. 965–78

Bourban P-E, Bögli A, Bonjour F, Månson J-A E 1998 *Compos. Sci. Technol.* **85** (5), 633–7

Charrier J-M 1990 *Polymer Materials and Processing*. Hanser Verlag, Munich, Germany

Cogswell F N (ed.) 1992 *Thermoplastic Aromatic Polymer Composites*. Butterworth-Heineman, Oxford

Månson J-A E 1994 In: Chawla K K (ed.) *New Demands on Composite Materials*. TMS, Warrendale, PA

Månson J-A E, Bourban P-E, Bonjour F 1995 *Pat. PCT*, WO 96/36477

Månson J-A E, Wakeman M D, Bernet N 2000 Composite processing and manufacturing—an overview. In: Kelly A, Zweben C (eds.) *Comprehensive Composite Materials*. Pergamon, Oxford, Vol. 2, pp. 577–608

Pang K P, Gillham J K 1990 *J. Appl. Polym. Sci.* **39**, 909–33

Schmid E, Eder R 1997 *German Pat.* 19 602 684 C1

C. J. G. Plummer, P.-E. Bourban and J.-A. E. Månson
Ecole Polytechnique Fédérale de Lausanne,
Switzerland

Polymer Melt Mixing: Agglomerate Dispersion

It is well known that the addition of fillers to polymeric materials can greatly improve their physical properties (see also *Elastomers, Reinforcement of**). Many of the additives used in the plastics and rubber industry are particulate in nature and exist in the form of aggregates composed of primary particles permanently fused together. In turn, aggregates cluster together under the influence of surface forces in collections of aggregates, referred to as agglomerates. A typical example is carbon black, with aggregates on the order of 0.1 μm in size and agglomerates in the range of 10–100 μm and up. The purpose of the mixing process is to break up the agglomerates into smaller size fragments and uniformly distribute these throughout the polymeric matrix. The process of dispersing particulate solids in a molten polymer involves a number of sequential, although most of the time overlapping, steps. These are: incorporation, deagglomeration, and distribution. In the incorporation step, the melt wets the particle surfaces, and the two phases are combined into a single coherent mass. At the end of this step, the loose filler particles disappear, and part of the air, introduced in the compound by entrapment with the agglomerates, has been replaced by the polymeric media. The polymer wets the filler, and it may penetrate into the void spaces of the agglomerate and also within the aggregates. In the deagglomeration, or dispersive mixing step, agglomerates or clusters of particles are broken down in size, either into their constituent particles or into smaller agglomerates. It is generally recognized that this is the most difficult step, and therefore the rate-determining one in any mixing operation. Once the filler agglomerates have been broken, separation of the closely spaced fragments and their distribution throughout the polymeric matrix should be accomplished.

In general, in polymer melt mixing operations, it is desirable, and for many applications it is essential, to eliminate large agglomerates, and therefore dispersion is regarded as the critical step in the mixing process. In order to investigate dispersion mechanisms, numerous experimental and theoretical studies of the breakup of agglomerates have been carried out. Bolen and Colwell (1958) were the first to propose that agglomerates rupture when the hydrodynamic drag forces acting on the surface of the cluster induce internal stresses which exceed a certain threshold value. Many investigators extended this basic idea and tried to develop models for the mixing process. Most of these models concentrated on the breakup of agglomerates into relatively large fragments. In general, however, it has been observed experimentally that the breakup of agglomerates can proceed by two distinct mechanisms (Rwei *et al.* 1990). One is the aforementioned rupture model, whereas in the other mechanism, erosion, individual particles or small clusters of particles are eroded from the surface of the parent agglomerate.

Any analysis of the agglomerate dispersion process must consider both forces involved in agglomerate breakup: namely, the forces of cohesion within the clusters and the hydrodynamic forces inducing their breakup. In the following we will briefly discuss these forces separately.

1. Cohesive Forces

Most of the present models for agglomerate cohesive strength treat the structure of the agglomerate in a very simplistic manner, as either a uniform porous sphere or as a uniform assembly of spheres, and link the strength of the agglomerate to interparticle forces. One example is the model developed by Rumpf (1962). For equal sized spherical particles assembled into an agglomerate, Rumpf derived an expression of the form:

$$T = \frac{9}{8}\frac{1-\varepsilon}{\varepsilon}\frac{H}{d^2} \tag{1}$$

where T is the tensile strength, d is the diameter of the spherical particles, ε is the agglomerate porosity, and H is the mean interparticle force. The model is based on certain important assumptions: (i) there are a large number of bonds in the fracture section; (ii) the bonds are randomly distributed over the fracture section and have random orientation in space; (iii) the particles in the agglomerate are randomly distributed in space; and (iv) all interparticle forces can be approximated by a single mean effective force.

A more complex model of agglomerate structure was developed by Sonntag and Russell (1987), who built agglomerates of polystyrene latex spheres as porous spheres with a position-dependent internal structure. The internal volume fraction within the floc, ϕ (r), at a local position r with respect to the center of mass, varies as r^{D-3}, where D is the mass fractal dimension of the floc. However, effects of fluctuations in volume fraction due to the presence of structural heterogeneities were ignored. Later on, Horwatt and co-authors (1992b) developed a method to analyze the influence of structural heterogeneities on agglomerate cohesivity. The authors simulated a number of different fractal and nonfractal structures. They searched through the different structures and identified the fragment(s) requiring the lowest stress to break. Such fragments were designated as critical fragments, determining the overall cohesivity of the agglomerate.

The interparticle forces depend on the mechanisms of adhesion between particles. Considering van der Waals interactions only, and making use of the Hamaker theory for the interaction potential between two spheres (Hamaker 1937), the expression for the mean interparticle force can be approximated by:

$$H = \frac{Ad}{24z^2} \tag{2}$$

where A is the Hamaker constant and z is the effective separation distance between particles. However, agglomerate porous structure may allow for the surface-driven infiltration of the surrounding polymeric matrix into the cluster. The infiltration of fluid can cause some rearrangement of the packing structure, or alter the cohesive forces binding the agglomerate together. Schubert *et al.* (1975) describe how liquid bridges within agglomerate may lead to increased cohesivity, due to the presence of capillary forces. For partially infiltrated agglomerates, the cohesivity can be calculated from:

$$T = S\,\frac{1-\varepsilon}{\varepsilon}\,\frac{6\gamma\cos\theta}{d} \tag{3}$$

where γ is the fluid surface tension, θ is the contact angle and S is the degree of saturation within the agglomerate.

2. *Hydrodynamic Forces*

In developing an analysis of agglomerate dispersion by an applied flow field, one has to take into consideration the effect of the applied hydrodynamics on the mechanics of the agglomerate. If agglomerates are viewed as homogeneous, impermeable spheres, undergoing bulk motion as single entities, the hydrodynamic stress acting on them can be derived using concepts originally suggested by Bagster and Tomi (1974). For linear flows of low Reynolds number, the hydrodynamic stress acting on the agglomerate surface, $\boldsymbol{\tau}$, can be written in the form:

$$\boldsymbol{\tau} = 5\mu(\mathbf{E}\cdot\boldsymbol{r}) \tag{4}$$

where $\mathbf{E}$ is the rate of strain tensor for the bulk flow field, μ is the viscosity of the polymer melt and $\boldsymbol{r}$ is the orientation vector relative to the sphere center. The approach taken to examine the stresses within the sphere is to consider any plane that divides the sphere into two parts. The fluid surface forces induce a total force on each of the two parts, from which a net hydrodynamic force on the dividing plane can be computed.

An extension of this model, accounting for the flow inside the porous agglomerate, was given by Adler and Mills (1979). The hydrodynamics is governed by the Stokes equation in the bulk fluid:

$$\nabla p = \mu\nabla^2\boldsymbol{u} \tag{5}$$

and the Brinkman equation within the porous agglomerate:

$$\nabla p = \mu\nabla^2\boldsymbol{u} - \frac{\mu}{\kappa}(\boldsymbol{u}-\boldsymbol{u}_\text{s}) \tag{6}$$

In the above equations, p is the pressure, $\mathbf{u}$ is the fluid velocity, κ is agglomerate permeability and $\mathbf{u}_\text{s}$ denotes the cluster velocity. The equations can be solved with boundary conditions of velocity and stress vector continuity on the surface of the sphere. To calculate the stresses inside the porous agglomerate, a condition of mechanical equilibrium in the solid can be applied:

$$\nabla\cdot\boldsymbol{\sigma} - \frac{\mu}{\kappa}(\boldsymbol{u}-\boldsymbol{u}_\text{s}) = 0 \tag{7}$$

where $\boldsymbol{\sigma}$ is the solid stress tensor. The second term represents the body force transmitted to the solid by the percolating fluid. More recently, Levresse and co-authors (1998), extended the analysis of stress distribution within partially-infiltrated agglomerates subjected to controlled flow conditions.

3. *Models for Agglomerate Dispersion*

In developing models of agglomerate dispersion, one must consider the agglomerate structure and cohesivity, as well as the effect of the applied hydrodynamics on the mechanics of the agglomerate. Either simple or more complex models can be used to describe the hydrodynamics and the cohesiveness of the agglomerate and to predict the locus of breakage.

Combining Rumpf's model for the cohesive strength of an agglomerate with a simple hydrodynamic analysis of two touching rigid spheres (suggesting the two halves of an agglomerate), Manas-Zloczower *et al.* 1982, 1985) derived a model for agglomerate dispersion in simple shear flows. The fraction of agglomerates broken by the applied flow field was found to be dependent on a dimensionless parameter scaling the strength of the flow field relative to the cohesive strength of the agglomerate. Agglomerate breakage was modeled as a repetitive process, with cleavage to occur always at the midplane, where the effect of hydrodynamic tension is largest.

Horwatt and co-authors (1992a, b), extended Bagster and Tomi's hydrodynamic analysis to study the influence of agglomerate structural heterogeneity on its dispersion behavior. Critical conditions for dispersion and ultimate fragment size distributions were found to be dependent on the connectivity of the agglomerate, and the intensity of applied stress relative to the strength of the connecting bonds.

The kinetics of agglomerate dispersion and the dynamics of particle size distribution will depend on the mechanism of dispersion i.e., erosion or rupture. In turn, the dispersion mechanism is controlled by system characteristics (clusters size and cohesivity, polymer melt viscosity, and interfacial properties), as well as by the geometry and strength of the applied flow field. In addition, in industrial polymer melt mixing equipment, conditions of agglomerate dispersion are space-and time-dependent. Thus, a comprehensive model for agglomerate dispersion has to consider the kinetics of the dispersion process in combination with a detailed analysis of the mechanics of flow in the mixing equipment.

See also: Polymer Melt Mixing: Distributive Mixing#; Polymer Melt Mixing: Liquid–Liquid Mixing#; Polymer Melt Mixing: Modeling in 3D#; Polymer Melt Mixing in Static Mixers#

Bibliography

Adler P M, Mills P M 1979 Motion and rupture of a porous sphere in a linear flow field. *J. Rheol.* **23**, 25–37

Bagster D F, Tomi D 1974 The stresses within a sphere in simple flow fields. *Chem. Eng. Sci.* **29**, 1773–83

Bolen W R, Colwell R E 1958 Intensive mixing. *Soc. Plast. Eng. J.* **14**, 24–8

Hamaker H C 1937 The London–van der Waals attraction between spherical particles. *Physica* **4**, 1058–71

Horwatt S W, Feke D L, Manas-Zloczower I 1992b The influence of structural heterogeneities on the cohesivity and breakup of agglomerates in simple shear flows. *Powder Technol.* **72**, 113–19

Horwatt S W, Manas-Zloczower I, Feke D L 1992a Dispersion behavior of heterogeneous agglomerates at supercritical stresses. *Chem. Eng. Sci.* **47**, 1849–55

Levresse P, Feke D L, Manas-Zloczower I 1998 Behavior of permeable agglomerates in prototype flow fields—relevance to agglomerate dispersion. *PPS North American Meeting Proc.*, pp. 86–7

Manas-Zloczower I, Nir A, Tadmor Z 1982 Dispersive mixing in internal mixers—a theoretical model based on agglomerate rupture. *Rubber Chem. Technol.* **55**, 1250–85

Manas-Zloczower I, Nir A, Tadmor Z 1985 Dispersive mixing in roll-mills. *Polymer Compos.* **6**, 222–31

Rumpf H 1962 *Agglomeration*. Wiley, New York, pp. 379–418

Rwei S P, Manas-Zloczower I, Feke D L 1990 Observation of carbon black agglomerate dispersion in simple shear flows. *Polym. Eng. Sci.* **30**, 701–6

Sonntag R C, Russell W B 1987 Structure and breakup of flocs subjected to fluid stresses II: theory. *J. Coll. Int. Sci.* **115**, 378–89

Schubert H, Herrmann W, Rumpf H 1975 Deformation behaviour of agglomerates under tensile Stress. *Powder Technol.* **11**, 121–31

I. Manas-Zloczower
Case Western Reserve University, Cleveland, Ohio, USA

Polymer Modified Wood

Polymer modified woods (PMWs) are kiln-dried, permeable lumber (such as birch, beech, red oak, maple, pine, poplar, rubberwood, and alder) or composites (such as plywood, oriented strandboard, particleboard, and fiberboard) that contain polymer (plastic) in their cell cavities or cell walls. PMWs are different from wood–plastic composites, which are largely a blend of thermoplastic resins and wood particles.

1. How PMWs Are Made

PMWs are made by impregnating the wood with liquid monomer or prepolymer mixtures and then polymerizing the liquid within the wood. Many monomers that contain vinyl groups (a carbon–carbon double bond) have a viscosity useful for wood impregnation at room temperature and also react well to form the polymer within the wood. The monomer most widely used for PMWs is methyl methacrylate.

Impregnation is carried out in a pressure vessel. First, a vacuum is drawn to remove air from the cells of the wood. Next, the liquid monomer is admitted to the pressure vessel while maintaining the vacuum. Atmospheric pressure is then restored to the vessel and additional pressure may be applied to help move the liquid uniformly into the wood. After impregnation the wood is removed from the vessel and placed in a curing chamber.

Curing (polymerization) of vinyl-type monomers starts with the generation of reactive groups using either temperature-sensitive catalysts or penetrating radiation such as gamma rays. The reaction proceeds with monomer molecules bonding together to form large polymer molecules and releasing heat as a by-product.

2. Types

There are three major types of PMWs. The cell-lumen type contains polymer, which fills cell cavities but has not reacted with the cell wall. Most vinyl monomers produce this type of PMW. Cell-wall PMWs are made using monomers, such as furfuryl alcohol monomer and phenol–formaldehyde prepolymer, which modify the cell wall. Phenol–formaldehyde prepolymer can be impregnated into thin veneers because of its high viscosity, and the impregnated veneers are glued together to make solid boards called “Impreg” and “Compreg.” Cell-wall PMWs have high dimensional stability, weatherability, and resistance to chemicals but are more brittle than the wood from which they are made. Combination monomer formulations, the third type, are designed to make PMWs with properties intermediate between those of cell-lumen and cell-wall PMWs.

PMWs are typically made by forcing as much liquid monomer into wood as possible, using what is called in the wood preservation industry the "full cell" process. Made this way, cell-lumen PMWs have a density very close to $1.1\,\mathrm{g\,cm^{-3}}$ and cell-wall PMWs $\sim 1\,\mathrm{g\,cm^{-3}}$ regardless of the wood from which they were made. Because of this uniform density, PMWs have more uniform piece-to-piece properties than wood.

2.1 *Cell-lumen PMWs*

The hollow cavities in wood cells makes them susceptible to crushing in side-loading and buckling in longitudinal compressive loading. In cell-lumen PMWs the cavities are filled or nearly filled with polymer. The individual cells are thus reinforced by the polymer in their cavities, dramatically increasing hardness and compressive strength of the wood perpendicular to the grain. Since buckling of cells is the first failure mode in bending, the reinforcing effect of filled cell lumens in these PMWs is noticeable in bending-mode mechanical values. Toughness is also increased because the polymer adds more material to the wood which must be fractured.

The blocking of cell cavities in cell-lumen PMWs greatly retards moisture movement, lowering water vapor diffusion coefficient values, and causes longitudinal and transverse movements to be about equal. Blockage of cell cavities should have a retarding effect on wood-boring insects and marine organisms. Overall swelling in room-temperature water is lowered. Since moisture plays a strong role in weathering, PMWs are more resistant to weathering than wood. As cell cavities are the main route of fungal invasion, filling them with polymer greatly retards decay.

Dyes can be added to the monomer mixture to make colored PMWs, which have the same physical properties, except for color, as those made with clear polymer.

2.2 *Cell-wall PMWs*

The wood cell walls in cell-wall PMWs are chemically altered. Since the polymer in cell-wall PMWs also partly fills cell lumens, cell-wall PMWs have some of the characteristics of cell-lumen PMWs. The most noticeably changed properties of cell-wall PMWs are those which depend on moisture movement and uptake, including swelling, water vapor diffusion coefficient, and weathering. Wood-destroying organisms do not degrade them and they are highly resistant to acids and bases. However, cell-wall PMWs are more brittle than the parent wood cell wall, which is their chief disadvantage. The dark-brown color of furfuryl alcohol-based PMWs is useful for dark-colored products, but the darkness sometimes limits their usefulness.

2.3 *Combination PMWs*

Combination PMWs attempt to obtain some of the better properties of both cell-lumen and cell-wall PMW with less of their respective disadvantages. In moisture effects, toughness, and decay resistance they are intermediate.

2.4 *Low Polymer Content PMW*

There is a class of wood treatment that causes a chemical reaction to occur in the cell wall. These processes use either a small amount of chemical or no chemical and heat. Such material is called "modified wood." Examples that have reached the commercial stage are acetylation and heat treatment.

It has been recently shown that cell-wall PMW with low polymer content has reduced shrinkage and swelling and is resistant to microorganisms. A small amount of monomer or prepolymer reacts with the cell wall to provide protection. Weight gains as low as 15% have been shown to provide increased lifetime in adverse environments. Change in mechanical properties is small because there is so little polymer added. This class of modified wood falls between PMW produced by a full cell monomer impregnation process and modified wood that uses small amounts of chemical or heat. One such process that is being commercialized is called "furfurylation" because it is based on polymerized furfuryl alcohol. Another process under development uses phenol-formaldehyde prepolymers.

Low polymer contents are obtained by using a solution of the monomer or prepolymer in a solvent to impregnate wood with a full cell impregnation process and then causing the solvent to evaporate. Water is the preferred solvent. For furfurylated wood, polymerization can be accomplished by heating before or during drying.

3. *Uses of PMWs*

Cell-lumen PMWs made from vinyl monomers have been used for many years in flooring, stair parts, and sporting goods. It is called "acrylic-impregnated wood flooring" in industry because it is produced using methyl methacrylate monomer. There is interest in using alternative monomer formulations and curing methods in other products, particularly where hardness, wearability, corrosion resistance, weatherability, decay resistance, moisture resistance, stiffness, color choice, and nontoxicity are desired. Examples include railway ties, highway posts, outdoor furniture, and counter-tops. When sanded and polished, PMWs have a deep luster superior in performance and appearance to most surface coatings. Adding color increases style options. The attractiveness of PMW products has also generated interest

Table 1
Examples of Brinell hardness values for wood and polymer-modified wood (PMW).

	Wood	Cell-lumen PMW	Cell-wall PMW
Sugar maple (*Acer saccharum*)	6	18	
Yellow birch (*Betula alleghaniensis*)	4	18	10
Scots pine sapwood (*Pinus sylvestris*)	2	12	

Table 2
Examples of physical properties of maple wood and PMW.

	Wood	Cell-lumen PMW	Cell-wall PMW
Percent weight loss from fungal decay in soil block test	93	13	0
24 h, 21°C percent tangential swell in liquid water	14	4	1
Longitudinal water vapor diffusion coefficient, $10^{-12}\,\mathrm{m^2\,s^{-1}}$	25	5	2
Transverse water vapor diffusion coefficient, $10^{-12}\,\mathrm{m^2\,s^{-1}}$	7	5	0.5

among craftspeople who make musical instruments, jewelry, and kitchenware.

"Impreg" and "Compreg" have been used for many years in knife handles, pistol grips, and other small items. With dyes added to the veneer layers, an attractive color pattern can be produced. Furfuryl alcohol cell-wall PMWs have been used for knife handles and kitchenware, which are exposed to frequent

Table 4
Property changes for low polymer content furfurylated wood ($\sim$30% weight gain based on wood). The variability in the furfurylated wood results from variations in original wood density and polymer content.

Properties	Untreated wood	Furfurylated wood
Physical		
Density (g/cc)	0.4–0.7	15–35% increase
Color		Slightly browner
Shrinkage		50–75% less
Hygroscopicity		50–75% less
Durability		
For use	Indoors and dry outdoors	Exposed outdoors and ground contact
Mechanical		
Compression parallel to grain		Same
Impact bending		20% less
Hardness (Brinell)	2–6	3–7

Table 3
Examples of mechanical properties (in MPa) of wood and cell-lumen PMW at oven-dry and 12% moisture content (MC).

	% MC	Wood	PMW
Sugar maple bending elastic modulus	0	14800	17200
	12	13100	15700
White pine bending elastic modulus	0	10300	14100
	12	6700	8900
Sugar maple tension parallel elastic modulus	0	16500	18000
	12	14600	16500
Sugar maple compression parallel elastic modulus	0	13300	16400
	12	11400	15100
Yellow birch compression perpendicular elastic modulus	0	1700	2900
Yellow birch compression perpendicular yield stress (elastic limit)	0	20	70
Yellow birch compression perpendicular ultimate (maximum) stress	0		120

mechanical dishwashing. It has potential as a substitute for some dark-colored tropical hardwoods.

Low polymer content PMW is similar in density, mechanical, and handling properties to untreated wood. Therefore, its potential is uses in which wood would be susceptible to decay and attack by insects and marine borers but a high polymer loading would be undesirable or unneeded.

4. Example Properties

Tables 1–4 give the properties of some PMWs and the woods from which they were made.

5. Conclusions

Filling wood with polymers produces a denser material with some properties noticeably enhanced over those of the parent wood. Therefore, making PMWs increases utility for some products for which wood is already used and also increases the applications in which wood can be used. Using low amounts of polymer to modify wood increases its moisture resistance and durability with little density increase or change to other properties.

See also: Wood: Macroscopic Anatomy*.

Bibliography

Furuno T, Imamura H, Kajita H 2004 The modification of wood by treatment with low molecular weight phenol-formaldehyde resin. *Wood Sci Technol* **37**, 349–61

Kumar S 1994 Chemical modification of wood. *Wood Fiber Sci.* **26** (2), 270–80

Meyer J W 1984 Wood–polymer materials. In: Rowell R M (ed.) *Chemistry of Solid Wood*, Advances in Chemistry Series. 207. American Chemical Society, Washington, DC, pp. 257–90

Rowell R M, Konkol P 1987 *Treatments That Enhance Physical Properties of Wood*, Report FPL-GTR-55, US Department of Agriculture, Forest Service, Forest Products Laboratory

Schneider M H 1994 Wood–polymer composites. *Wood Fiber Sci.* **26** (1), 142–51

Westin M, Lande S, Schneider M 2003 Furfurylation of wood—process, properties and commercial production. *Proceedings of European Conference on Wood Modification*

M. H. Schneider
University of New Brunswick, Fredericton, NB, Canada

Polymer Nanocomposites with Metal Dispersions

Polymer composite materials incorporating finely divided metal particles in the nanometer size regime combine a number of attractive aspects. Especially intriguing are the unique properties which can arise from the nanosized metal particles ("*size effects*"), resulting in materials possessing novel catalytic, magnetic, optical, or electronic properties. The polymeric matrix can contribute additional properties, giving, for instance, processibility and/or film-forming qualities. Furthermore, an improved protection from surrounding media (such as air or water vapor), or the prevention of agglomeration and precipitation of the metal particles, is provided.

1. Preparation Methods

The conventional way of obtaining polymer–metal composites involves mechanical routes, such as mixing, blending, or milling of the polymers and the metal components. Another approach involves the generation of the metal nanoparticles *in situ*, that is, in the presence of the polymeric matrix. In this case, well-dispersed metal particles in the nanometer size regime can be obtained. Before the generation of the particles the metal precursor has to be homogeneously incorporated into the polymeric matrix. This can be achieved in solution, by combining diluted solutions of the polymer and the metal precursor prior to reduction in solution (the *dispersion route*). Alternatively, the combined solutions containing the polymer and the metal precursor can be deposited onto a substrate, and the reactions (in most cases thermal or photochemical decompositions) are then performed within the solid film (the *deposition route*). A third approach involves the immersion of the solid polymer in a solution containing the metal precursor and allowing the precursor to diffuse into the swollen polymeric matrix (the *immersion route*). The three preparation routes for the metal–polymer nanocomposite materials are depicted schematically in Fig. 1.

The zerovalent metal colloids are frequently obtained by various wet-chemistry reductions. A large number of reducing agents have been investigated, common examples being tetrahydroborates, hydrazine, or formaldehyde. Nanoparticle features, specifically particle size, size distribution, shape, and morphology, can be strongly influenced by selection of the reducing agent. This is mainly due to the large effect it can have on particle nucleation and growth.

Other preparation methods include thermal decompositions or photolyses of labile metal precursors, photoreductions of light-sensitive precursors, such as silver and gold salts, or evaporation methods in a dispersing medium. The use of high-intensity ultrasound, e.g., for the generation of iron particles in the presence of poly(*N*-vinyl-2-pyrrolidone), has been demonstrated as well (Suslick *et al.* 1996).

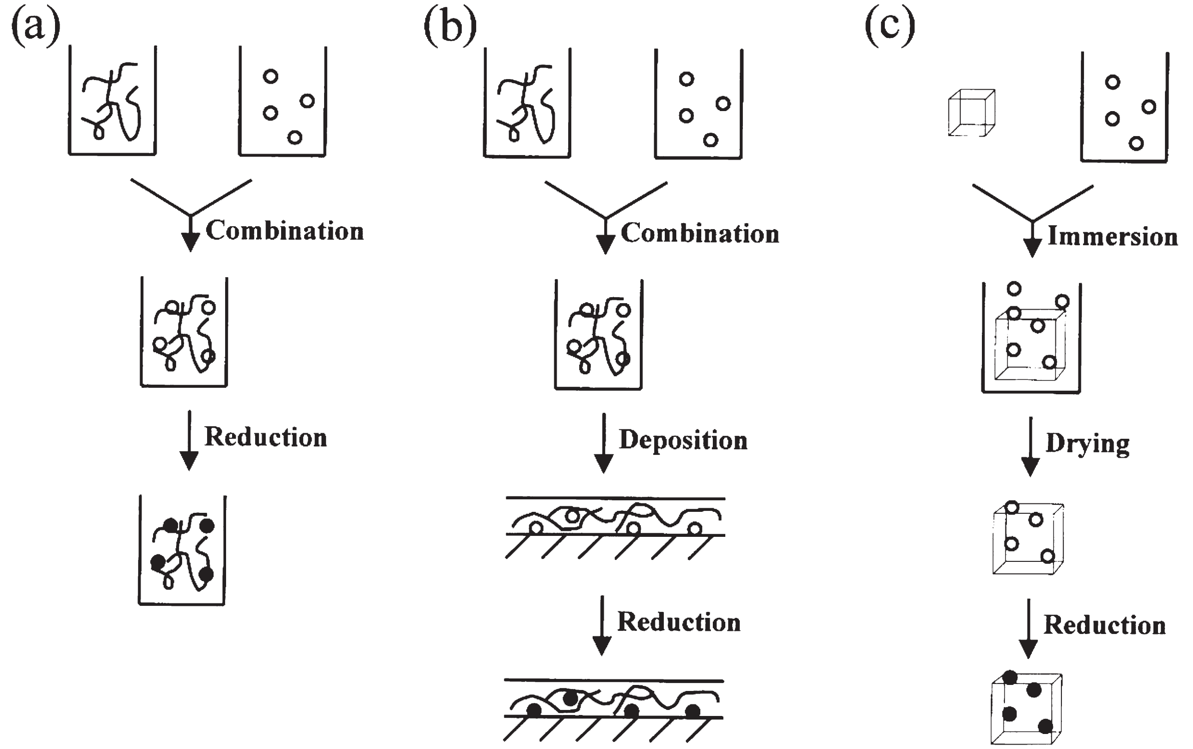


Figure 1
Preparation methods for polymer–metal nanocomposite materials: (a) dispersion route, (b) deposition route, and (c) immersion route.

2. *Polymeric Matrices*

Frequently, the polymers are mainly employed as the dispersing and stabilizing media for the metal nanoparticles (*protective polymers*). However, in many cases they can also influence the particle and materials properties directly, and thus provide a means to "fine tune" and tailor the materials for specific applications (*functional polymers*). For instance, since the polymer is present during colloid preparation, it can exert a direct influence on particle nucleation and growth. Such effects are usually due to interactions between certain polymer components and the metal species which are possible during various stages of preparation. For instance, complex or ion pair formation with the metal precursor species can lead to small particle sizes and narrow size distributions.

Figure 2 shows an example of a transmission electron micrograph, of platinum particles generated in the presence of the cationic polyelectrolyte poly(3-chloro-2-hydroxypropyl-2-methacryloxyethyl-dimethyl ammonium chloride) by reduction with potassium tetrahydroborate. In this case the negatively charged metal precursor species hexachloroplatinate, $[PtCl_6]^{2-}$, can form ion pairs with the positively charged ammonium side-groups of the polyelectrolyte. This results in the homogeneous distribution of the metal precursor and control of the nucleation process.

During reduction and particle growth it is also possible that the polymer or its components adsorb on the growing nanoparticle. With preferred adsorption along certain crystal planes this can result in the control of the nanoparticle shapes, by directing the particle growth in certain directions. For instance, the use of various polyacids as protective agents frequently leads to the formation of single-crystal features for gold nanoparticles generated in their presence by a mild reduction method. An example is shown in the transmission electron micrograph of Fig. 3, for gold nanoparticles obtained in the presence of poly(vinyl phosphonic acid) by photoreduction. Several gold nanoparticles exhibiting triangular profiles can be recognized, in addition to fewer hexagonal and pentagonal ones.

After reduction the polymer is responsible for the long-term colloidal stabilization of the metal colloids, which occurs frequently by hydrophobic interactions between hydrophobic segments of the polymer and the surfaces of the metal nanoparticles.

Figure 2
Platinum nanoparticles generated from a dihydrogen hexachloroplatinate (H_2PtCl_6) precursor by reduction with potassium tetrahydroborate in the presence of poly(3-chloro-2-hydroxypropyl-2-methacryloxyethyldimethyl ammonium chloride). Small platinum nanoparticles exhibiting a narrow size distribution are obtained (bar = 22 nm).

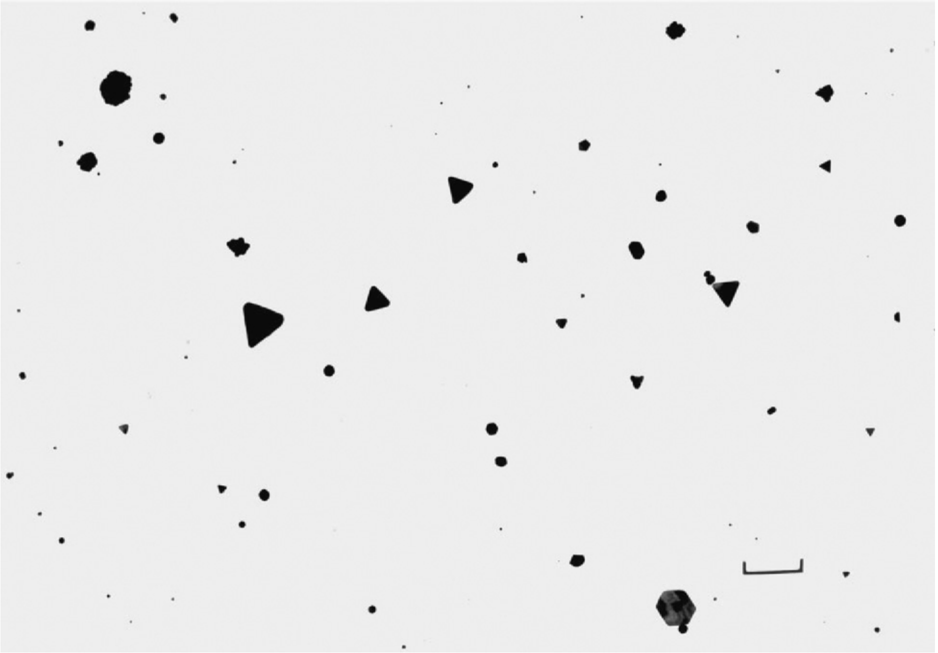


Figure 3
Gold nanoparticles generated from a chloroauric acid ($HAuCl_4$) precursor by photoreduction in the presence of poly(vinyl phosphonic acid). Several triangular particle profiles indicating the formation of single crystals can be recognized, in addition to a few hexagonal and pentagonal ones (bar = 245 nm).

A large variety of polymeric matrices are of interest, depending on the applications and the desired properties of the materials. Therefore, the types of polymer investigated cover a broad spectrum, ranging from nonionic homo- and random copolymers, to polyelectrolytes, to block and graft copolymers. The stabilization of the metal nanoparticles is usually based on *steric stabilization* (Bradley 1994), schematically shown in Fig. 4. Steric stabilization, both by homo- and random copolymers, and by block copolymers, is based on two effects: (i) the reduced configurational freedom of the polymer chains in the region between particles causes the lowering of entropy (*entropic effect*), and (ii) the increased concentration of the polymer in the interparticle region results in an increased local activity which is counteracted by solvation (*osmotic effect*) (Bradley 1994). Especially effective is the *electrosteric stabilization* provided by the use of charged polymeric species. In this case the effects for steric stabilization are combined with the electrostatic repulsion stemming from the charged polymers and associated counterions.

The polymers should ideally act as good protective agents for the stabilization of colloidal metal nanoparticles. A quantitative measure of the *protective function* of such agents was first introduced by Zsigmondy as the *gold number* (Zsigmondy 1901). The gold number is defined as the amount of the protective colloid in milligrams which just prevents 10 cm^3 of a red gold sol from changing color to violet upon addition of 1 cm^3 of a 10% aqueous solution of sodium chloride.

Especially when employing polar or aqueous solvents, polymers possessing a relatively hydrophobic backbone (which can interact with the surface of the metal colloids), and hydrophilic side-groups (which can interact with the polar medium and provide the solubility) have been found most suitable. Good protective abilities have been reported for such polymeric species (Hirai and Toshima 1986). Flexible, water-soluble vinyl polymers, such as poly(*N*-vinyl-2-pyrrolidone) or poly(vinyl alcohol) are currently the most extensively studied representatives. For organic solvent media or thin-film applications, chemically stable and heat-resistant polymers such as poly(methyl methacrylate) or poly(vinylidene fluoride) have been employed.

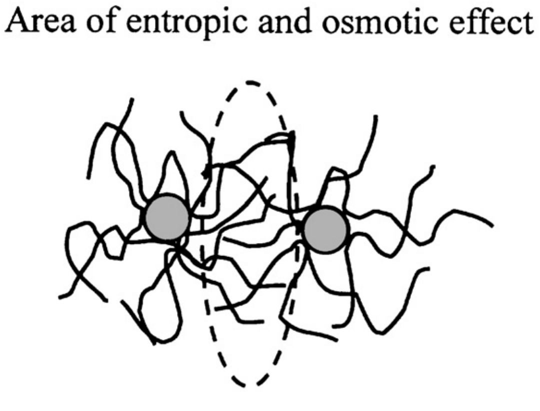


Figure 4
Steric stabilization of metal nanoparticles by polymeric protective agents.

Block copolymers are of particular interest. Due to the capability of this group of polymers for phase separation and formation of various morphologies, highly versatile composite materials can be developed. Especially promising is the use of spherical micelles as "*nanoreactors*" for the generation of metal nanoparticles (Bronstein *et al.* 1998). This involves the uptake of the metal precursor into the micelle core before reduction, which is usually achieved by complex or ion pair formation. This results in the localization of nanoparticle generation in or near the micelle cores, and can furthermore result in a tremendous size control of the metal nanoparticles, due to the spatial confinement created by the micelles (Bronstein *et al.* 1998).

3. Catalytic Properties

It is anticipated that nanometer-sized metal particles dispersed within a polymeric matrix will be especially useful for the development of tailored catalytic materials. Apart from novel catalytic properties and the possibility of tuning them, these materials are expected to provide additional advantages, such as facilitated handling and storage, and improved recoverability. Due to the small particle sizes, very large surface areas are provided, resulting in highly active catalyst systems.

The materials can be employed in various ways. A considerable amount of work has been performed using colloidal solutions for liquid phase catalytic reactions (Bradley 1994). However, the use as solids or gels is equally intriguing, and such materials have been investigated for catalytic gas phase reactions (Ciebien *et al.* 1998).

Noble metal particles protected by synthetic polymers and their use as catalysts have been already reported by Dunworth and Nord (1954) for the liquid phase hydrogenation of a large variety of unsaturated organic compounds, ranging from nitrobenzenes to unsaturated oils.

Several different catalytic reactions and materials have been explored. Most extensively studied are liquid phase hydrogenations of olefins, cycloalkenes, or acetylenes involving palladium, platinum, or rhodium nanoparticles. As expected, a dependence of the catalytic activity on the particle size has been found; however, the type of polymer is equally important for influencing the catalytic properties.

Selectivity in such catalytic reactions has been demonstrated as well, mainly using water-soluble polymers, such as poly(*N*-vinyl-2-pyrrolidone), in liquid phase reactions. It has been generally found that the rate of hydrogenation depends on the structure of the substrates. Thus, the hydrogenation of terminal double bonds (e.g. 1-hexene) is usually preferred to cyclic (for instance, cyclohexene) and internal (e.g., 2-hexene) ones.

Similarly, the selective hydrogenation of acetylenes to olefins, and of dienes to monoenes, has been performed. It has been proposed that the selectivity is achieved by the steric and hydrophobic layer created by the polymer, which surrounds the metal particle and interacts with its surface. Apart from the diffusion of the reactant through the polymeric layer toward the catalyst surface, a certain geometry on the catalyst surface is required for a catalytic reaction to occur. Steric hindrance provided by adsorbed polymeric species prevents such necessary geometrical arrangement of the substrate, especially for certain hindered positions of the double bonds, or for larger reactants. Similarly, the steric hindrance provided by the polymer also decreases the mobility and possible conformations of the reactant close to the catalyst surface. This can be important for selective hydrogenations involving multiple double bonds or triple bonds. The selective hydrogenation of dienes to the corresponding monoenes was furthermore explained by the preferred adsorption of the dienes on the nanocatalyst surface, due to the small particle sizes involved. For particles below about 5 nm a larger number of surface atoms possessing a coordinatively unsaturated state are present, resulting in preferred adsorption of reactants possessing the larger number of double bonds (Hirai *et al.* 1985).

Selective catalytic reactions were also reported for solid film catalyst systems, which could be incorporated into a membrane (Ciebien *et al.* 1998). In these cases gas phase catalytic reactions were examined, involving a mixture of reactants. High catalytic activities and selectivities have been found for the gas phase hydrogenation of 1,3-butadiene, and for a mixture of ethylene and propylene (with the preferred hydrogenation of ethylene).

An option for the design of *intelligent catalyst systems* is provided by differences in the solubility of the polymeric matrix. Chen and Akashi (1997) have shown that variations in the catalytic activities are observed with the change of the solubility of the protective polymer in a respective solvent. With the use of poly(*N*-isopropylacrylamide) a temperature-sensitive polymeric matrix was chosen, which allowed the control of the rate for the liquid phase hydrogenation of allyl alcohol to 1-propanol, simply by changing the temperature of the reaction mixture. The observed differences in catalytic activity are explained by more or less strong adsorption of the polymer on the catalyst surface, due to changes in solubility of the polymer at different temperatures.

The investigations include not only monometallic systems but bimetallic systems as well. These materials are usually prepared by the simultaneous *in situ* reduction of a mixture of the metal precursors in the desired ratio. In this way, catalyst systems containing platinum–rhodium, rhodium–gold, palladium–platinum, platinum–gold, or palladium–gold colloids have been obtained. Usually these bimetallic catalyst nanoparticles exhibit core–shell structures, which have a profound influence on the catalytic properties of the systems. (Yonezawa and Toshima 1995). The catalytic activities and selectivities of these bimetallic systems are often increased in comparison to the respective monometallic catalyst systems, and also depend on the metal ratio and composition.

Especially interesting are results on the application of metal–polymer systems for the Heck reaction (Klingelhöfer *et al.* 1997), which demonstrate the utility of such metal–polymer systems for a large variety of catalytic reactions. The findings have shown the effectiveness of palladium colloids in block copolymer micelles for catalyzing the carbon–carbon coupling of aryl halides with alkenes (Heck reaction). A dependence of the catalytic properties on the size of the block copolymer micelles has been found, which allows further catalyst modifications by the choice of the polymer block lengths and types.

The catalytic properties of polymer-protected noble metal nanoparticles are furthermore significant for the evolution of hydrogen from water induced by visible light. This photochemical process offers a very attractive possibility for the conversion of solar energy into chemical energy, and therefore is of profound technological importance (Grätzel 1981, Bradley 1994). Usually, this process involves multicomponent systems containing (i) a photosensitizer which is excited by visible light, (ii) an electron relay system (consisting of an electron donor and an electron mediator), and (iii) a redox catalyst. The redox catalyst system is involved in the catalytic reduction of water, resulting in the generation of hydrogen. Stabilized noble metal colloids, in particular platinum colloids or bimetallic species (such as gold–platinum), have been found especially efficient as redox catalyst systems.

4. Magnetic Properties

Hybrid materials incorporating particles possessing magnetic properties are very promising for a variety of applications, including information storage/magnetic recording, ferrofluids, various biomedical applications (such as diagnostic systems), or the improvement of the mechanical properties of materials.

The most thoroughly investigated metals are iron, nickel, and cobalt; however, a larger amount of work is available on magnetic materials involving metal oxides, such as iron oxides (Tannenbaum *et al.* 1985).

A common feature of these less-noble zerovalent metal colloids is their sensitivity toward oxidation in an ambient atmosphere. Therefore, their preparation is mostly performed in an inert atmosphere, and the oxidation of the outer layer of the particles occurs in most cases immediately upon contact with air. Usually, a core–shell structure is formed, possessing a metal core and a thin oxide layer. The oxidation process can be followed by the change of the magnetic properties of the materials (Griffiths *et al.* 1979).

The most prominent property of nanometer-sized ferromagnetic or ferrimagnetic metal particles is their *superparamagnetic behavior*, which is observed below a critical size of the particles (usually about 10 nm). Therefore these systems are also interesting as models of small magnetic particles. Below the critical size it is energetically favorable for a particle to contain a single magnetic domain. For such small particles the magnetic energy becomes comparable to the thermal energy, and the typical bulk magnetic characteristics, such as retentivity, coercivity and hysteresis, at room temperature disappear. Upon decrease of the temperature such superparamagnetic particles reach a *blocking temperature* at very low temperatures where their magnetic moment is fixed, and a hysteresis in the magnetization curve, due to magnetic anisotropy, appears (Ziolo *et al.* 1992, Sohn and Cohen 1997).

A general observation is made for fluids, such as colloidal solutions of nanometer-sized magnetic particles. They exhibit no anisotropic effects, and superparamagnetic behavior is usually found. This is caused by the ability of the colloidal metal particles to rotate and align under the application of a magnetic field. For sufficiently large particle sizes the magnetic dipole–dipole interaction energy becomes larger than the thermal energy, and a tendency of the particles to form linear chains is observed (Thomas 1966). For instance, for cobalt, such chain formation starts at particle sizes of about 10 nm. In colloidal solutions or at temperatures above melting of a polymeric matrix, the particles have sufficient mobility to allow chain formation to proceed to a high degree.

Such behavior has been employed for the alignment of magnetic metal nanoparticles dispersed in polymeric matrices hardened under the application of a magnetic field. This has caused formation of a "chain structure" of the ferromagnetic particles, resulting in anisotropic electrical conductivities of the metal-filled polymeric materials (Gul' and Golubeva 1967).

The incorporation of magnetic metal particles into polymers can furthermore be of interest for improving the mechanical properties of such composite materials. A variety of materials, mainly based on oxides, has been explored (Chow and Penwell 1986). Homogeneously dispersed palladium nanoparticles of about 10–20 Å diameter have been used as fillers in poly(methyl methacrylate), resulting in reinforcement and the increase of the elastic modulus (Nakano *et al.* 1993). It has been shown that the mechanical properties can increase remarkably with a decrease in the filler particle size. The incorporation of metallic species can also contribute to the cross-linking of the polymeric matrix. For instance, poly(vinylidene fluoride) has been cross-linked during the decomposition of iron carbonyls. This was apparent from a change in the polymer solubility properties and from mechanical measurements as well (Reich and Goldberg 1983).

The incorporation of metallic filler particles is interesting, especially when combined with magnetic properties, since these can influence the particle orientation and alignment under the application of a magnetic field. Furthermore, the application of a magnetic field can counter particle-settling due to gravity, which is especially important during solvent removal (Kusy 1986). In this way, highly homogeneous metal–polymer materials exhibiting improved mechanical properties could be obtained.

5. Summary

Polymer composite materials containing dispersed nanometer-sized metal particles can be readily prepared and are useful for a variety of applications. The unique properties of these materials, which mainly arise from the nanometer size dimensions, are promising not only for catalytically active and magnetic systems: the exploration of their optical and electronic properties is equally significant, and could lead to the development of intriguing novel materials. The presence of the polymeric matrix adds to the versatility by providing additional functions, such as the processibility or solubility of the composites. It can also profoundly influence the materials properties. This allows the development of polymer–metal nanocomposite materials which can be tailored for specific applications and needs.

Bibliography

Bradley J S 1994 In: Schmid G (ed.) *Clusters and Colloids. From Theory to Applications*. VCH, Weinheim, Germany

Bronstein L M, Valetsky P M, Antonietti M 1998 In: Fendler J (ed.) *Nanoparticles and Nanostructured Films. Preparation, Characterization and Applications*. Wiley-VCH, Weinheim, Germany

Chen C-W, Akashi M 1997 Preparation of ultrafine platinum particles protected by poly(*N*-isopropylacrylamide) and their catalytic activity in the hydrogenation of allyl alcohol. *J. Polym. Sci. Part A: Polym. Chem.* **35**, 1329–32

Chow T S, Penwell R C 1986 In: Bhattacharya S K (ed.) *Metal-filled Polymers. Properties and Applications*. Dekker, New York

Ciebien J F, Cohen R E, Duran A 1998 Catalytic properties of palladium nanoclusters synthesized within diblock copolymer films: hydrogenation of ethylene and propylene. *Supramol. Sci.* **5**, 31–9

Dunworth W P, Nord F F 1954 Noble metal-synthetic polymer catalysis and studies on the mechanism of their action. *Adv. Catal.* **6**, 125–41

Grätzel M 1981 Artificial photosynthesis: water cleavage into hydrogen and oxygen by visible light. *Acc. Chem. Res.* **14**, 376–84

Griffiths C H, O'Horo M P, Smith T W 1979 The structure, magnetic characterization, and oxidation of colloidal iron dispersions. *J. Appl. Phys.* **50**, 7108–15

Gul' V E, Golubeva M G 1967 Investigation of electric conductive anisotropic structures in polymer materials. *Koll. Z.* **29**, 62–8

Hirai H, Chawanya H, Toshima N 1985 Colloidal palladium protected with poly(*N*-vinyl-2-pyrrolidone) for selective hydrogenation of cyclopentadiene. *React. Polym.* **3**, 127–41

Hirai H, Toshima N 1986 In: Iwasawa Y (ed.) *Tailored Metal Catalysts*. Reidel, Dordrecht

Klingelhöfer S, Heitz W, Greiner A, Oestreich S, Förster S, Antonietti M 1997 Preparation of palladium colloids in block copolymer micelles and their use for the catalysis of the heck reaction. *J. Am. Chem. Soc.* **119**, 10116–20

Kusy R P 1986 In: Bhattacharya S K (ed.) *Metal-filled Polymers. Properties and Applications*. Dekker, New York

Nakano N, Hasegawa S, Nakao Y 1993 Effect of reinforcement of elastic modulus in ultra-microscopic particle filled polymer composite materials. *J. Soc. Mater. Sci., Jpn.* **42**, 1072–6

Reich S, Goldberg E P 1983 Poly(vinylidene fluoride)-γ-Fe_2O_3 magnetic composites. *J. Polym. Sci., Polym. Phys.* **21**, 869–79

Sohn B H, Cohen R E 1997 Processible optically transparent block copolymer films containing superparamagnetic iron oxide nanoclusters. *Chem. Mater.* **9**, 264–9

Suslick K S, Fang M, Hyeon T 1996 Sonochemical synthesis of iron colloids. *J. Am. Chem. Soc.* **118**, 11960–1

Tannenbaum R, Goldberg E P, Flenniken C L 1985. In: Sheats J E, Carraher Jr C E, Pittman Jr C U (eds.) *Metal-Containing Polymeric Systems*. Plenum Press, New York

Thomas J R 1966 Preparation and magnetic properties of colloidal cobalt particles. *J. Appl. Phys.* **37**, 2914–5

Yonezawa T, Toshima N 1995 Mechanistic consideration of formation of polymer-protected nanoscopic bimetallic clusters. *J. Chem. Soc. Faraday Trans.* **91**, 4111–9

Ziolo R F, Giannelis E P, Weinstein B A, O'Horo M P, Ganguly B N, Mehrotra V, Russell M W, Huffman D R 1992 Matrix-mediated synthesis of nanocrystalline γ-Fe_2O_3: a new optically transparent magnetic material. *Science* **257**, 219–23

Zsigmondy R 1901 Die Hochrothe Goldlösung als Reagens auf Colloide. *Z. Anal. Chem.* **40**, 697–719

A. B. R. Mayer
Polytechnic University Brooklyn, New York, USA

Polymer Transcrystallinity in Composites

Transcrystallinity (TC) in composites can give rise to some of the most spectacular microstructures found in materials science and technology. Two examples of transcrystalline structures around aramid fibers in polypropylene, a thick TC layer in an isothermally crystallized specimen and a fine TC region for a fast-cooled specimen, are shown in Fig. 1. Even though TC layers have well-defined morphologies, it is a matter of some debate as to whether or not such microstructures have any effect upon the resulting properties of the fiber-reinforced composite. The interest in TC stems from the possible effects it could have on the interface between fiber and matrix in fiber-reinforced composites. Could TC at the interface enhance fiber/matrix adhesion? Could it form a layer around the fiber in which the matrix properties are slightly different from the bulk? If this were the case could it protect the fiber from bulk residual stresses, or perhaps it could increase the stress transfer to the fiber from the matrix? This article first defines and describes TC and then addresses the effect it has upon the fiber/matrix interface and composite properties.

1. *Definition of Transcrystallinity*

TC can be defined as a restricted crystalline growth from many closely spaced heterogeneous nuclei. The growth front must be a line, plane, or smooth surface parallel to the substrate initiating the nuclei. It has been proposed that geometric considerations determine that for a series of nuclei to produce TC the distance between the nuclei should be one-fifth, or less, of the distance from the nuclei to the growth front (Schoolenberg and van Rooyen 1993).

Another restricted crystalline growth resulting from many closely spaced nuclei, but which is not always differentiated from TC (Thomason and van Rooyen 1992a), is the cylindrite (Varga and Karger-Kocsis 1993). Though sometimes optically indistinguishable from TC, cylindrites are defined as being nucleated homogeneously, or nucleated by the imposition of shear at the interface. Cylindrites are created in polypropylene by inducing shear between substrate and melt, thus the homogeneous nuclei from which the cylindrite grows may be nucleated heterogeneously on the substrate.

A third way in which to generate a restricted crystalline growth from many closely spaced nuclei is to use a chemical additive at the interface to nucleate the matrix (Varga and Karger-Kocsis 1993). No specific name has been ascribed to the resulting structure.

2. *Morphology*

2.1 *Transcrystallinity—Orientation*

TC in polymers is a result of a semicrystalline oriented growth process (Folkes 1995). The orientation of the lamellas within TC (Waddon *et al.* 1987) is important because some physical properties, such as thermal expansion coefficients or moduli, could depend on lamella orientation. Stern *et al.* (1997)

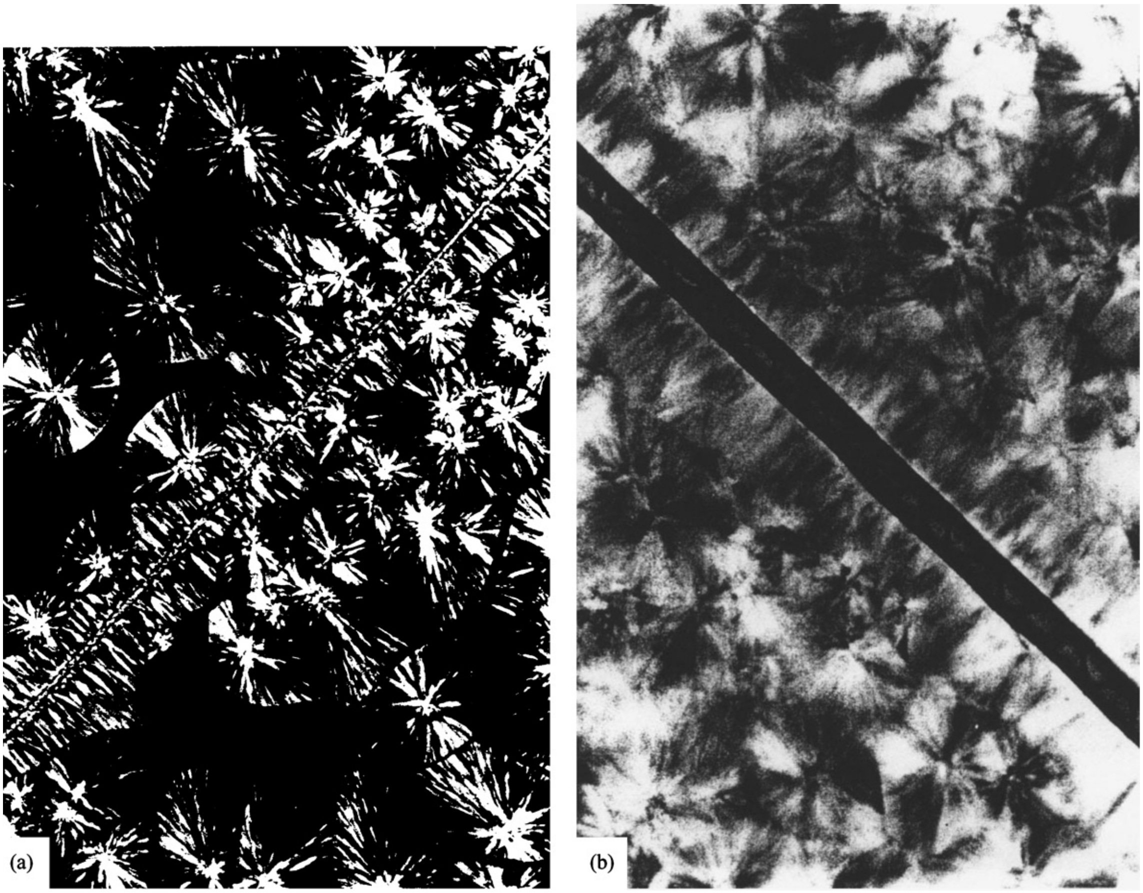


Figure 1
Optical micrographs of TC in an aramid/polypropylene single fiber model composite. (a) Specimen crystallized isothermally at 130 °C to give a thick transcrystalline layer, (b) specimen cooled rapidly to give a narrow transcrystalline layer (fiber diameter 12 μm).

used x-ray diffraction on polyethylene TC and found that the morphology of the TC was dependent on the thermal history and identical, except for orientation, to that of the surrounding spherulites.

2.2 Cylindrites

Not all publications differentiate between TC and cylindrites. Some call shear-induced TC "transcrystallinity" (Thomason and van Rooyen 1992a,1992b), some avoid calling it TC and some, notably those written by Varga and Karger-Kocsis (1993), make a point of calling naturally induced structures, TC and shear-induced structures, cylindrites.

2.3 Kinetics

During the process of analysis of TC, the growth rate of TC had to be compared with that of spherulites in comparable circumstances. It has been found that the linear growth rate of TC is the same as the radial spherulite growth rate under the same conditions (Waddon *et al.* 1987).

3. Factors Affecting the Production of Transcrystallinity

3.1 Thermal Conditions

It is found that the temperature, T_m, at which the matrix is melted, the crystallization temperature T_c, the cooling rate, and the differential contraction providing orientation, all contribute to the presence or lack of TC in a composite.

Avella *et al.* (1992) found that by increasing the melt temperature and time in melt, the nucleation density on Kevlar fibers decreases for polypropylene. They also showed that by increasing T_c for polypropylene the production of TC becomes less likely, and the interface becomes increasingly spherulitic. Thomason and van Rooyen (1992a) measured the growth rate of polypropylene TC with increasing T_c and found it to be comparable with the equivalent spherulite growth.

Faster cooling rates are found to be favorable to the formation of TC in polypropylene (Table 1). Thomason and van Rooyen (1992b) systematically compared the morphologies of single-fiber polypropylene composites produced using glass, carbon, and aramid fibers. Given a sufficiently high cooling rate (100 K min^{-1}) even glass fibers induced TC (see Table 1).

An increasing probability of TC with increased cooling rate is due to the shear in the melt against the fiber induced by their different thermal contractions during cooling. Thomason and van Rooyen (1992a, 1992b) tested this assertion by comparing the morphology of a wide variety of fiber/polypropylene single fiber composites at one given thermal history (Table 2). The ability to nucleate TC was compared to the axial thermal expansion coefficients of the fibers. For the given thermal history, all the fibers (four nylon, two aramid, and three high modulus carbon fiber) with axial thermal coefficients, α, less than $-0.4 \times 10^{-6}\,\mathrm{K}^{-1}$ produced TC. Fibers with larger axial thermal coefficients (high strength carbon, E-glass, and alumina) did not produce TC. The axial thermal coefficient of the polypropylene is $+300 \times 10^{-6}\,\mathrm{K}^{-1}$, thus it was confirmed that the greater difference caused TC.

3.2 Surface Topology

The nucleation of TC has often been ascribed to the surface topology or characteristics of the fiber. This could be due to the melt nucleating against the fiber because the unit cells match sufficiently well for epitaxial growth to occur (Chen and Hsaio 1992) or because the fiber is sufficiently rough for the melt to be mechanically caught and nucleate there (Avella *et al.* 1992). The wide variety of matrices and fibers in which TC can be found, however, implies that a similarity in the crystallographic unit cell of fiber and

Table 1
Effect of cooling rate upon the formation TC in polypropylene.[a]

Cooling rate (K min^{-1})	Glass	Carbon	Aramid
0.2		No	No
1		No	No
2		?	?
5	No	?	?
10	No	?	Yes
50	?	Yes	Yes
100	Yes	Yes	
280	Yes		
α ($\times 10^6$ K) axial	+4.9	−0.5	−6.8

Source: Thomason and van Rooyen (1992a).
[a] Yes/no indicates if a transcrystalline morphology was obtained.

Table 2
Fiber axial thermal expansion coefficients (After Thomason and van Rooyen 1992a, 1992b).

Fiber type	α ($\times 10^{-6}$ K) axial	Transcrystallizes polypropylene?
Polymeric		
ICI 1142 (nylon-66)	−185	Yes
ICI 1352 (nylon-66)	−160	Yes
Enka 155HRS (nylon-66)	−160	Yes
Enkalon 540T (nylon-6)	−130	Yes
Aramid		
Kevlar 49	−6.3	Yes
Twaron D1035	−3.5	Yes
Carbon high modulus (HM)		
Thornel P120	−1.44	Yes
Thornel T50	−0.9	Yes
Enka HM35	−0.5	Yes
Carbon high strength (HS)		
Thornel T40	−0.2	No
Enka ST III	−0.1	No
E-Glass	+4.9	No
Alumina	+15.0	No

matrix is not necessary for TC to occur. In fact, Thomason and van Rooyen (1992a) produced TC in isotactic polypropylene (iPP) on a glass fiber (Table 1), which is amorphous, thus showing that crystallinity in the substrate is not even necessary.

Nonuniform surface topology is also known to generate TC as it is thought that the molecules in the melt may become aligned in cracks on the fiber surface. Avella *et al.* (1992) also suggested that TC nucleated on unmelted fragments of polymer nuclei hidden in the surface flaws of the fiber surface. Chen and Hsiao (1992), on explaining why high-modulus (HM) carbon fibers nucleate TC in a number of aromatic polymers, whereas high-strength (HS) carbon fibers do not, proposed that the larger graphite planes in HM fibers were the determining factors in inducing TC. Thomason and van Rooyen (1992a) compared several HM and HS carbon fibers for their ability to produce TC in polypropylene (Table 2). They found, however, that HS carbon fibers have higher axial thermal expansion coefficients than HM ones and are thus less likely to produce TC.

3.3 Molecular Weight

As the molecular weight of a polypropylene melt is increased, the tendency to transcrystallize is reduced. Thomason and van Rooyen (1992a) studied the effect of molecular weight on TC in polypropylene and concluded that higher molecular weights allows TC to be created at higher crystallization temperatures. This is explained by the reduced ability of a longer chain melt molecule to relax from an oriented state caused by the difference in axial thermal expansion coefficients.

3.4 Other Factors

A number of other factors have been associated with the formation of TC and several have been discounted. For example, the surface energy of the substrate per melt pair does not appear to control the nucleating ability of the substrate (Chen and Hsiao 1992).

Factors which are known to affect the nucleation process include the commercial source of a polymer, and the presence and nature of a coating or matrix nucleant on the fiber or in the matrix (Chen and Hsiao 1992). Thomason and van Rooyen (1992a, 1992b), after considering thermal history, axial thermal expansion coefficients, and molecular weight, also noted that TC did not form over the whole length of shorter fibers (since the center would not shear the melt sufficiently).

Strain in a molten polymer can cause row nucleation in the bulk melt by orientation of the molecules. Shearing a substrate within a melt can also cause multiple nucleation at the interface followed by oriented restricted growth (which is often called TC).

4. Physical Properties

There is accumulated evidence that the Young's modulus of TC material is higher than that of the bulk. Klein *et al.* (1995) proved experimentally (using microtomed dynamic mechanical thermal analysis specimens) that this was so for nylon 6,6 TC by comparing TC and bulk polymer made under the same conditions. Folkes and Hardwick (1987) found that polypropylene TC had a greater Young's modulus and tensile yield strength than the spherulitic morphology and that the flexural modulus and strength of TC material was a little higher than that for the spherulitic samples. The two morphologies that were compared were prepared using different thermal histories, however, and it is known that higher crystallinity can produce a higher flexural modulus, strength, and fatigue life.

Teishev and Marom (1995) found that the transverse tensile strength of PE/PE composites was smaller than expected. From scanning electron microscope (SEM) analysis of the failed composites they attribute this to weaknesses present at TC growth front interfaces. They also found that the longitudinal tensile strength appeared to be unaffected by TC.

Tregub *et al.* (1994) found that a carbon fiber PEEK composite was stronger in the longitudinal direction than was expected. It was suggested that the TC present reduced the residual stresses, thus strengthening the composite. Klein *et al.* (1995) postulated that an alignment of crystallites would reduce the thermal expansivity mismatch between fiber and matrix, and thus reduce the thermal stresses exerted on the fiber.

5. Adhesion Properties of Transcrystalline Materials to Substrates

5.1 Residual Stresses

High levels of residual stress can be induced in the fibers when TC structures are prepared (Heppenstall-Butler *et al.* 1996). This is because the microstructures are prepared at elevated temperatures and cooled to room temperature. The large difference in thermal expansion coefficient between the fiber and matrix (Table 2) and the volume shrinkage of the matrix on crystallization, lead to residual compression in the fibers. This is particularly apparent in single-fiber model composite specimens (e.g., by kink band formation as seen in Fig. 1(b)) where the fibers may even fail in compression due to matrix shrinkage (Heppenstall-Butler *et al.* 1996). Hence, care must be taken in extrapolating the behavior of single-fiber model composites to that of high volume fraction systems.

5.2 Chemical Modification of the Interface

The effect of chemicals at the fiber/matrix interface upon TC has been found to depend on the chemicals and combination used. It is generally found that the interface is improved by the presence of silanes even though they can reduce the nucleating ability of the substrate and, hence, the appearance of the TC in the material. Without a coating the adhesion may be poor even when TC is present.

5.3 *Shear*

There have been a number of attempts to determine the shear strength of the interface in model composites with transcrystalline morphologies. The behavior depends upon the fiber/matrix system under consideration but, in general, large improvements in interfacial strength are not found with the presence of TC. The best way to determine the strength of the interface is to prepare single fiber TC pull-out specimens and then measure the stress required to pull the fiber out. Figure 2 shows the behavior of a single aramid fiber being pulled out of a TC polypropylene matrix. The point-to-point variation of stress in the fiber was determined from stress-induced Raman band shifts (Heppenstall-Butler *et al.* 1996). The strength of the interface in the aramid/polypropylene system was found to be about 10 MPa and independent of the presence or absence of TC.

5.4 *Failure Modes in Samples with Transcrystallinity*

The position of failure in a sample containing TC should indicate the weakest constituent of the composite which may be at one of four positions as indicated in Fig. 3. Waddon *et al.* (1987) found that failure in carbon-fiber/PPS films occurs between the TC and the spherulites rather than inside the TC due to boundary structural weakness (Fig. 3, position 3). In contrast, Tregub *et al.* (1994), using multi-fiber carbon-fiber/PEEK composites found that fracture took place through the TC region (Fig. 3, position 2).

Yue and Cheung (1993) presented SEMs of glass fibers after pull-out from polypropylene. These showed lumps of nuclei still adhered to the fiber surrounded by a fine array of radiating lines. From this they concluded that the nuclei adhered better to the fiber than the rest of the spherulite did, and thus suggested that TC would adhere better than a spherulitic morphology since there would be more nuclei at the surface.

Another TC failure process obtained after single fiber fragmentation was found by Wagner *et al.* (1993). They describe failure by *treeing*, a process where a crack propagates from the fiber perpendicularly up through the TC layer and continues into the spherulites beyond following the radial lamellas. On reaching the center the crack splays out down various radial paths (Fig. 3).

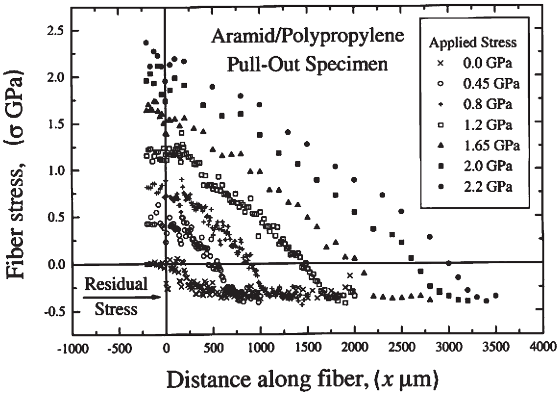


Figure 2
Pull-out of an aramid fiber from a polypropylene matrix, showing variation of fiber stress with distance along the fiber (position where the fiber enters the matrix, $x = 0$) at different level of applied stress. The specimen was cooled rapidly to give a microstructure similar to that shown in Fig 1(b). The level of residual stress in the fiber is indicated.

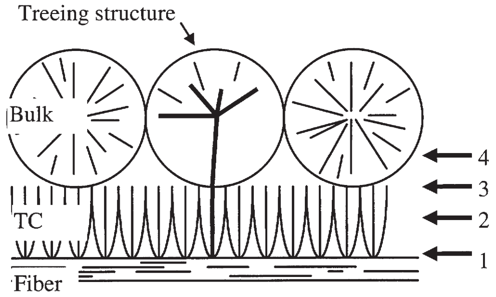


Figure 3
Schematic representations of possible positions of failure. (1) At the fiber/matrix interface, (2) in the TC material, (3) at the TC/spherulitic interface, and (4) in the bulk spherulitic material.

6. *Concluding Remarks*

To form TC in any fiber/matrix pair it is not necessary for the fiber to be crystalline. The formation of TC in a specific fiber/matrix pair is facilitated by large fiber/matrix thermal expansion coefficient mismatches, high cooling rates, and is molecular weight dependent. Inducing shear in the melt by pulling the fiber always results in a TC-like structure that is often referred to as a *cylindrite*.

TC has been found to possess a morphology similar to that of the bulk of the sample, but with different local molecular and lamellar orientation. The thermal expansion coefficient of TC has been found to vary parallel and perpendicular to the fiber. It is also found that the elastic modulus of TC samples is higher than that of the bulk whereas the viscoelastic damping modulus is lower than that in the bulk. The physical properties of the matrices are found to vary with thermal history.

Most of the investigations undertaken so far into the effect of TC upon composites have employed single-fiber model specimens. The presence of TC in such test-pieces is found to have only a marginal effect upon the strength of the fiber/matrix interface. It remains to be seen if the findings upon single-fiber model composite specimens can be translated to high volume fraction composites where interfiber distances are finite and the behavior is dominated by fiber rather than matrix properties.

Bibliography

Avella M, Della Volpe G, Martuscelli E, Raimo M 1992 Transcrystallinity phenomena in a polypropylene Kevlar fiber system. 1. Influence of crystallization conditions. *Polym. Eng. Sci.* **32**, 376–82

Chen E J H, Hsiao B S 1992 The effects of transcrystalline interphase in advanced polymer composites. *Polym. Eng. Sci.* **32**, 280–6

Folkes M J 1995 Interfacial crystallization of polypropylene in composites. In: Karger-Kocsis J (eds.) *Polypropylene Structure, Blends and Composities, Vol. 3 Composites*. Chapman and Hall, London, Chap. 10

Folkes M J, Hardwick S T 1987 Direct study of the structure and properties of transcrystalline layers. *J. Mater. Sci. Lett.* **6**, 656–8

Heppenstall-Butler M, Bannister D J, Young R J 1996 A study of transcrystalline polypropylene/single-aramid fibre pull-out behavior using Raman spectroscopy. *Composites* **27**, 833–8

Klein N, Marom G, Pegoretti A, Migliaresi C 1995 Determining the role of interfacial transcrystallinity in composite-materials by dynamic-mechanical thermal-analysis. *Composites* **26**, 707–12

Schoolenberg G E, van Rooyen A A 1993 Transcrystallinity in fiber-reinforced thermoplastic composites. *Compos. Interfaces* **1**, 243–52

Stern T, Marom G, Wachtel E 1997 Origin, morphology and crystallography of transcrystallinity in polyethylene-based single-polymer composites. *Composites* **28**, 437–44

Teishev A, Marom G 1995 The effect of transcrystallinity on the transverse mechanical-properties of single-polymer polyethylene composites. *J. Appl. Polym. Sci.* **56**, 959–66

Thomason J L, van Rooyen A A 1992a Transcrystallized interphase in thermoplastic composites. 1. Influence of fiber type and crystallization temperature. *J. Mater. Sci.* **27**, 889–96

Thomason J L, van Rooyen A A 1992bb Transcrystallized interphase in thermoplastic composites. 2. Influence of interfacial stress, cooling rate, fiber properties and polymer molecular-weight. *J. Mater. Sci.* **27**, 897–907

Tregub A, Harel H, Marom G 1994 Thermal-treatment effects on the crystallinity and the mechanical-behavior of carbon-fiber poly(ether ether ketone) composites. *J. Mater. Sci. Lett.* **13**, 329–31

Varga J, Karger-Kocsis J 1993 The occurrence of transcrystallization or row-nucleated cylindritic crystallization as a result of shearing in a glass-fiber-reinforced polypropylene. *Compos. Sci. Technol.* **48**, 191–8

Waddon A J, Hill M J, Keller A, Blundell D J 1987 On the crystal texture of linear polyaryls (PEEK, PEK and PPS). *J. Mater. Sci.* **22**, 1773–84

Wagner H D, Lustiger A, Marzinsky C N , Mueller R R 1993 Interlamellar failure at transcrystalline interfaces in glass/polypropylene composites. *Compos. Sci. Technol.* **48**, 181–4

Yue C Y , Cheung W L 1993 Some observations on the role of transcrystalline interphase on the interfacial strength of thermoplastic composites. *J. Mater. Sci. Lett.* **12**, 1092–4

R. J. Young and M. Heppenstall-Butler
University of Manchester Institute of Science and Technology, UK

Polymer-based Nanocomposites by Sol–Gel Routes, Applications of

The fabrication of inorganic nanoparticles by chemical means has become an interesting route to prepare many kinds of nanostructured materials. For example, sol–gel techniques, which have been used for over 60 years for making inorganic materials, can be considered as an interesting basis for the synthesis of nanoparticles, since sols as colloidal systems in general are composed of stabilized nanoparticles in solution. With a few exceptions, for example, SiO_2 sols prepared from alkoxide precursors at low pH-values, most sols are of particulate nature. These sols can be not only prepared from alkoxide precursors through the sol–gel process but through many precipitation reactions carried out outside of the point of zero charge. In these regimes, depending on pH and the chemical nature of the surface, particles absorb ions and develop surface charges which prevent aggregation, precipitation, or gel formation. In Fig. 1, some principles for the stabilization of sols are shown. Whereas in aqueous solutions most of the precipitates carry surface groupings, for example OH-groups, which tend to react with each other by the formation of chemical or hydrogen bonds, the presence of surface modifiers which do not have groupings to form

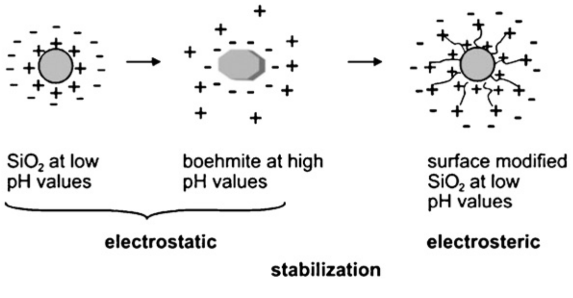

SiO₂ at low pH values
boehmite at high pH values
surface modified SiO₂ at low pH values
electrostatic
electrosteric
stabilization

Figure 1
Various types of stabilization of colloids.

strong bonds can prevent irreversible aggregation even at the point of zero charge. In Fig. 1, several types of surface modification are shown. The electrostatic stabilization is very effective in diluted sols and outside of the point of zero charge. In this case sols, stable over months, can be prepared. In the case of electrosteric stabilization, a surface is covered with groupings not able to form strong bonds to the next colloid, so that only van der Waals or weak polar forces are effective. In this case the aggregation leads to weak aggregates, which can be redispersed under appropriate conditions. The formation of strong and weak bonds is schematically shown in Fig. 2.

The preparation of surface-modified colloids is not only of interest for stabilization, but also for generating the desired reactivities of the colloids. In Fig. 3 several possibilities are shown. Acids (1) and bases (4) can be used to tailor an appropriate zeta-potential or to carry out condensation reactions to anhydrides, acid chlorides or acids. (2) and (3) (epoxides or methacrylates), represent polymerizable or polycondensable surface groupings which can be used for further reaction, for example, coupling reactions, but also for polymerizing the nanoparticles together with organic monomers in order to fabricate so-called polymer matrix nanocomposites. The alkyl group (5) represents a class of nonreactive groupings which can be used for reducing the particle-to-particle interaction, and also to adapt polar nanoparticles to nonpolar matrices. The combination of these surface-modifying principles with the inorganic nanoparticles' properties leads to a wide range of possible new composite materials. Nanoparticles by themselves have interesting properties, as shown in Tables 1 and 2. In Table 1, basic properties mainly resulting from the particle size are shown. One of the very interesting properties is the large surface area, especially together with matrix materials, since interfaces, for example, between inorganic surfaces and organic polymers in general have a structure different from the polymer matrix structure.

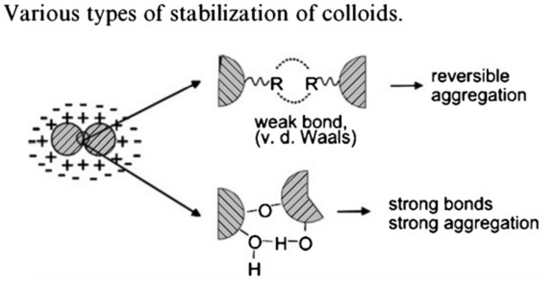


Figure 2
Aggregation mechanism for the formation of "hard."

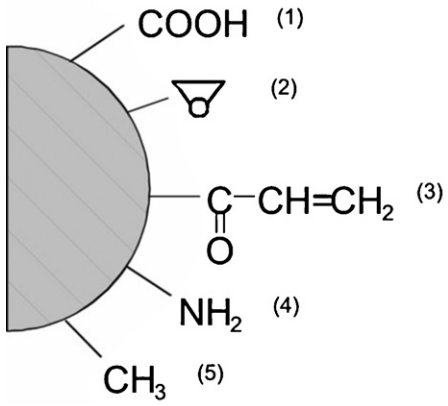


Figure 3
Various types of functional surface modification.

In Table 2, some examples for the relations between colloid types, functional groupings, materials, and basic properties are shown.

1. Fabrication of Nanoparticles

This is based on the rules of colloidal chemistry in connection with surface modifying agents, which at the same time act as thermodynamic stabilizers, and which not only prevent the reaction of the particles with each other, but also keep the particle size at a desired level during the colloid formation process (for example, precipitation). This approach has been successfully used for making various kinds of nanoparticles (Schmidt 1996, Schmidt and Nonninger 1998). This is schematically shown in Fig. 4. The figure also shows that by using small organic molecules, the volume fraction of the coating is rather low compared to the nanoparticle size. If the modifying agent is bifunctional, that means that not only is it linked to the surface by chemical bonds, as shown in Fig. 4 for complex bonds on oxidic particles by β-diketones, but the surface modification leads to the desired reactivities on the particle surface. In addition to this, and as also shown by Schmidt (1996) and Schmidt and Nonninger (1998), the desired crystallinity of the

Table 1
Properties of nanoparticles depending on their small size

Surface area	interfacial phase catalytic properties sensor properties
Size effects	superparamagnetic or superferromagnetic properties
Quantum effects	semiconducting plasmon resonating NLO photocatalytic effects
Highly dispersed nanocomposites	abrasion resistance + transparency
Transparency	refractive index

Table 2
Examples of nanoparticle effects in nanostructured materials

Particles or chemical functions	Type of material	Properties/application
Ceramic nanoparticles (e.g., SiO_2, Al_2O_3, ZrO_2)	polymer- or hybrid matrix-nanocomposites	refractive index: interference layers lenses; GRIN-lenses scratch resistance: coatings ion conductivity: smart windows
Ceramic or functional nanoparticles + functional groups (CF_3, …loxy, epoxy, amino acids, etc.)	hybrid matrix nanocomposites; polymer matrix; nanocomposites with colloids linked to the matrix by polymerization	low surface energy coatings low refractive index coatings hard coatings smart coatings corrosion protection coatings
Metal colloids	glass matrix	colors; non-linear optics
Transition metal colloids	ceramics, glass matrix	colors, catalysts, sensors
Semiconductors (e.g., TiO_2, CdS)	coatings, glass matrix, hybrid matrix nanocomposites	refractive index tailored coatings photocatalytic coatings nonlinear optics holograms
Luminescent nanoparticles	ceramic coatings, polymer matrix	transparent luminescent security systems transparent chromophores
Transition metal nanoparticles	inorganic coatings	catalytic systems

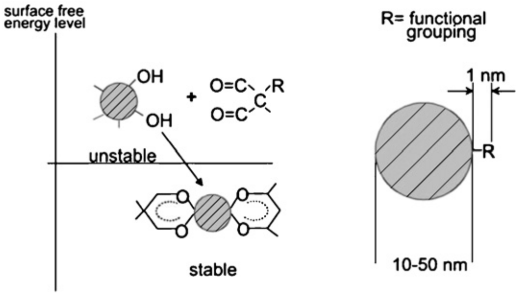


Figure 4
Thermodynamic effect of surface modification.

particles can be obtained by subsequent hydrothermal treatment. This means that using the chemical approach with appropriate surface modifiers, nanoparticles with desired surface reactivities can be produced and used as a basis for the nanocomposite fabrication. If these particles are combined with or dispersed in appropriate matrices, for example, in polymers, organic monomers or other hybrid sol–gel precursors such as organic alkoxy silanes, nanocomposites can be fabricated by many different routes. If the nanoparticles are produced within a monomeric sol–gel or polymer precursor system, this is called the *in situ* route. If the nanoparticles are produced separately, isolated and then redispersed in appropriate matrix systems, it is called the *separate processing* route. These nanoparticles have also been introduced into a polymeric matrix by compounding processes, and these processes have been already successfully applied to transparent coatings or bulk materials (R. Nonninger: personal communication, 1998).

In the following section, a series of examples is given showing how the nanoparticle route can be used for various material developments.

2. Optical Applications

2.1 Smart Materials (*Photochromics, Electrochromics, Luminescence*)

So-called smart materials are causing more and more industrial interest for many reasons. Photochromic materials are very common in the field of bulk glass, e.g., for ophthalmic glasses. A new approach in this field is photochromic nanocomposite hard coatings as wet coating systems, using photochromic dyes. The photochromic effect of common dyes like spirooxazines or spiropyrans is based on a change of the conformation, which changes the size of the organic π-bond systems and hence the electron energy levels of the dye molecule. With an increase in size of the π-system, the energy levels for the excitation of the electrons decrease in number. This means that the absorption may be shifted from the UV into the

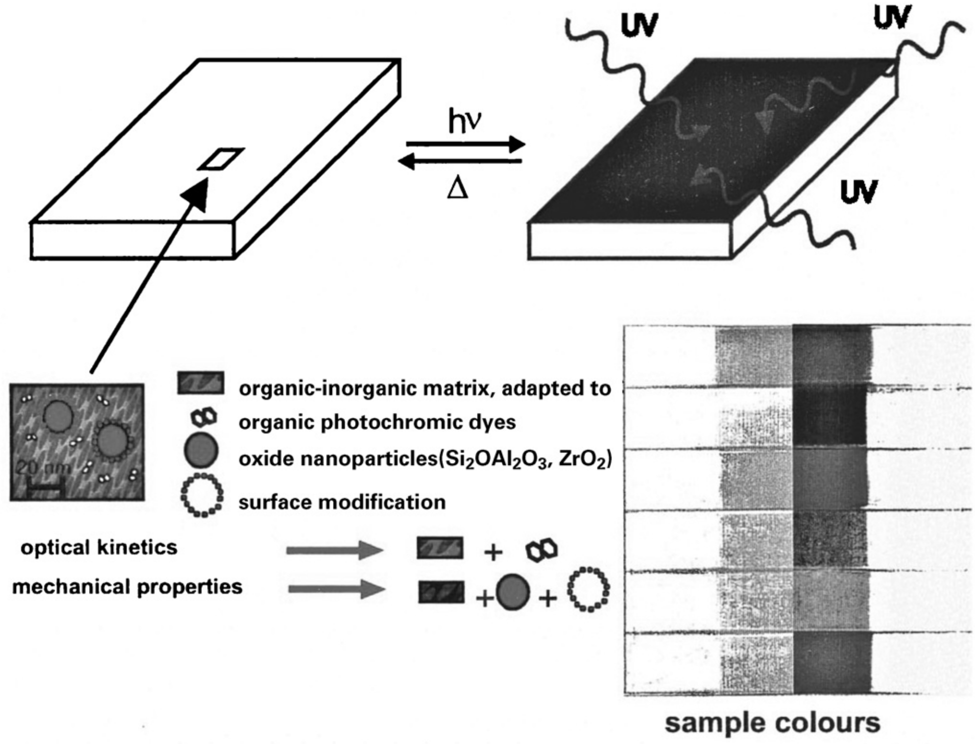

UV
UV
UV
hν
Δ
20 nm
organic-inorganic matrix, adapted to
organic photochromic dyes
oxide nanoparticles(Si2OAl2O3, ZrO2)
surface modification
optical kinetics
mechanical properties
sample colours

Figure 5
Photochromic coatings on glass without and with UV-irradiation (after Hou *et al.* 1994).

visible. This conformation change requires a soft "host matrix system," such as polymers with sufficient free volume. The more rigid the matrix system is, the more difficult it is for the molecule to change conformation. In order to use photochromic dyes in coatings, however, a high abrasion resistance is required. This in general can be obtained only by a very rigid matrix. In order to overcome these contradictory requirements, a nanocomposite coating material has been developed with ceramic nanoparticles—for example, from alumina or zirconia—which, when well dispersed in the matrix so as to obtain a good scratch resistance (where a soft matrix system made from polyacrylic components is synthesized in the voids between the nanoparticles), now provides sufficient flexibility for the photochromic dyes. The aim of the development is a photochromic coating material based on organic photochromic dyes with response times of seconds, which can be applied on transparent substrates (glass, polymers) or on nontransparent substrates (ceramics, polymers, paper) for the preparation of coatings with a thickness of 5–50 µm. The method is based on the incorporation of spirooxazines, pyrans, and fulgids into the matrix of an organic–inorganic composite, which is prepared by a specially developed sol–gel process as described elsewhere (Fig. 5; Hou *et al.* 1994).

The result of this development is that by using UV radiation, a darkening of from 90% to 10% transmission can be reached. By further modification of the components of the matrix, the kinetic behavior of the photochromic dye can be adapted to different applications: glasses, pane glass, plastic foils (transparent and opaque), paper, ceramics, etc. The half-life times for darkening and fading ($t_{0.5}$) can be established anywhere between 2 s and 10 min. Due to the modularity of the network structure, different photochromic dyes can be integrated, resulting in neutral tints (gray or brown).

One of the superior features of nanoparticles and transparent matrices is that Raleigh scattering can be neglected, if particle size and refractive index differences are appropriate. This is shown in Eqn. (1):

$$\gamma_{\text{ext}} = 32\pi^4 c \frac{n_p^2 - n_0^2\, r^3}{n_p^2 + 2n_0^2 \lambda^4} \tag{1}$$

where n_{p} is the refractive index, n_0 is the reflective index of the matrix, r is the particle radius, γ is the scattering loss, λ is the wavelength, and c is a constant. From this equation a general rule can be deduced, which shows that if the particle size is below 1/20 of the wavelength, Raleigh scattering becomes

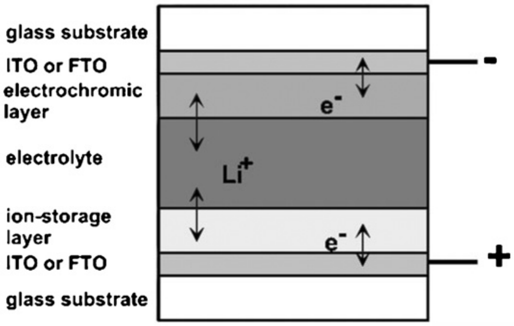


Figure 6
Schematic layout of an electrochromic (EC) device.

very low and can be neglected in most cases, at least in thin films. From this point of view, nanocomposites or nanoparticulate thin films can be used for optical purposes. The second very interesting feature is that the physical, optical, or electronic properties connected to the crystalline or inorganic structure of a nanoparticle are still present, and can be used for obtaining specific properties.

A very interesting example of nanoparticle technology is smart windows or electrochromic windows, based on tungsten oxide. In Fig. 6, the multilayer system of an electrochromic cell is shown. The glass substrates are coated with transparent electronic layers, made from for example ITO or FTO, which are in direct contact with the electrochromic layer, which is made from tungsten oxide, vanadium oxide, molybdenum oxide, niobium oxide, or nickel oxide. Due to an electrochemical reduction, mixtures of different oxidation states can be obtained, leading to high absorption in the visible range. For compensating the charges, protons or lithium atoms have to be diffused into the electrochromic layers. On the other side of the cell a similar process takes place which, in general, does not lead to a change of the visible spectrum. In this case oxides like those of cerium or titanium can be used. This layer also has to show a certain storage capacity for the charge-compensating ions like protons or lithium. In order to avoid electronic shortcuts, an electronically isolating layer, having a high ionic conducting capacity, has to be used. For this reason, polymers with a sufficient solubility for lithium sols or acids are used (for example polyethylene oxide). These systems have been stabilized by the introduction of nanoparticles, for example zirconia. At their surfaces, lithium or protonic transport takes place advantageously. Moreover, the nanoparticles help to improve the mechanical properties of these materials, because at the same time they have to be used as adhesives for gluing the layered systems together. It was demonstrated (Munro *et al.* 1998) that if the conductivity of the lithium containing nanocomposites reaches the region of about 10^{-4} S, and the electronic conductivity in the ceria, titania, and tungsten oxide system can be kept low, the switching behavior of such cells based on nanotechnologies is almost independent of the thickness of the ionic conductor, and systems up to 1 mm in thickness can be used.

This is very important for the production of large areas, since the planarity of large glass panes is very difficult to establish at the micrometer level. The chemical nanotechnology allows the production of cheap devices, because it has become possible to produce large area wet coatings, of optical quality, with very reproducible film thicknesses (Mennig *et al.* 1998b). The reversible variation in transmission of such cells reaches 70% (550 nm, bleached state) down to 21–25% (coloured state), with switching times t_{80} (t_{80} = time taken for 80% of transmission change): of 120 s (coloration), 70 s (bleaching) for EC-cells of $50 \times 80\,\mathrm{cm}^2$. The benefits of this system are more comfort and energy saving in the architectural and automotive areas (e.g., air-conditioning costs in summer) and privacy effects.

By segmenting such an electrochromic device into pixels, and with an appropriate computer-controlled voltage supply, the fabrication of electrochromic displays becomes possible.

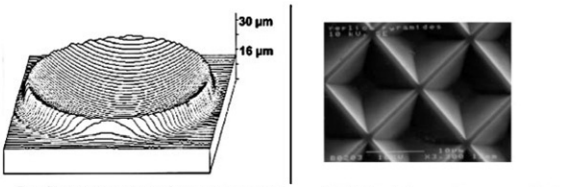
This opens up the possibility of the fabrication of large-area monochromatic displays between panes of glass, for example for displaying special sales offers in shop windows. Luminescent nanoparticles, such as those based on doped Eu^{3+}, have been developed, because there is a demand in industry for materials with high specific emission values. Interesting applications include photoluminescent displays, security tags on paper, plastic, luminescent textiles, etc. This development leads to materials in the form of redispersable nanoscaled powder, sols, coatings, and bulk materials, which present high luminescent emission under UV excitation. The main issue is the synthesis of rare-earth luminescent complexes (Eu^{3+}, Tb^{3+}, etc.), which perform the so-called antenna effect. These complexes can be incorporated into organic modified sols or attached on redispersable crystalline nanopowders. By this procedure, luminescent transparent coatings have been developed with thicknesses of up to 3 μm deposited on various glasses, polycarbonate, paper, or textiles.

2.2 Micropatterned Coatings

The fabrication of patterns in films is an interesting technology for many applications, such as the fabrication of holograms, the storage of information, the fabrication of microlens arrays, or even the microelectronic patterning and surface mounting device fabrication processes in microelectronics. Several basic techniques have to be distinguished. Etching techniques are very common in microelectronics. Patterning techniques using embossing require moldable systems, for example thermoplastic polymers. Embossing of sol–gel layers, as shown by Lukosz and Tiefenthaler (1983) and Tohge *et al.* (1994), requires very high pressures. In most cases, there are serious restrictions on using thermoplastic polymers, since it is difficult to obtain high precision patterns, due to the relaxation of the patterns as long as the polymer temperature is above the glass transition temperature T_g; because of the limited film thicknesses, only patterns with a very low step height are possible. This means that it is very difficult to produce precise structures from pure inorganic systems like SiO_2, TiO_2, or glasses. In addition to this, polymers have low mechanical resistivity—for example, scratch resistance is poor.

The micropatterning of inorganic–organic nanocomposites in combination with UV-curing led to a process which allows the fabrication of micropatterns in a hard coating type of material, with a UV-curing during the embossing step. The curing is based on the use of methacryloxysilane, which can be photo-cross-linked. Microlens arrays and refractive gratings have been developed using this approach. Another approach is based on the fact that SiO_2 nanoparticles coated with methyl groupings lead to gels with a remarkably increased relaxation behavior, and s film thicknesses of 20 μm and more are possible (Mennig *et al.* 1992). Due to the softness of these gels, they can be patterned with silicone rubber replicas and even curved shapes can be embossed. After embossing and densification at 500 °C, a transformation of the organically modified SiO_2 sol into almost pure SiO_2 glass (98% of the theoretical density) takes place. In Fig. 7(a), a profile of an embossed nanomer microlens is shown; in Fig. 7(b), a pyramidal structure of SiO_2 on glass as a light trap for the collection of light in solar collectors.

Through the application of a two-wave mixing process, holographic techniques have been developed. In this case, the intensity modulation of the incident light is used in a nanoparticle-containing film, with a nanoparticle surface modified with polymerizable groupings (Judeinstein *et al.* 1997). Before this, an inorganic backbone was built up in the system by hydrolysis and condensation, but care must be taken that no cross-linking of the nanoparticles to the system takes place. In the highly irradiated zones, the polymerizable nanoparticles are consumed by the polymerization process, and from the darker zones unpolymerized nanoparticles diffuse into the irradiated zones. By this technique an increase in concentration of nanoparticles in the irradiated zones takes place. If the nanoparticles show a higher refractive index than the matrix—for example, if zirconia or tantalum oxide nanoparticles are used—then volume phase holograms can be prepared. Meanwhile, the fabrication of the holograms by a two-way fixing process can be performed within seconds, due to the high diffusion rates of the nanoparticles in the matrix. After the holographic process is finished, a fixation of the matrix by full-area photopolymerization takes place. In Fig. 8, some samples of the fabrication of such gratings are shown (diffraction gratings, Fresnel lenses). Alternatively, instead of being fixed after the holographic process, the gratings can be developed by dissolving the unradiated areas. An EDX-pattern of a holographically fabricated refractive grating is also shown in Fig. 8. The concentration fluctuation of the tantalum oxide as a function of the two-wave mixing irradiation is clearly demonstrated.



Profilometer plot of an embossed microlens in a Nanomer film

(a)

HSEM picture of a pyramidal structure in a SiO_2 film

(b)

Figure 7
Lenses, produced by micropattering of SiO_2 films.

As a variation of these "moving nanoparticles," the nanoparticles also can be set into motion by electric fields, if the appropriate zeta potential is established in a non-fully-cross-linked matrix. Using this approach, gradient lenses up to 5 cm in diameter have been developed (Oliveira *et al.* 1996).

2.3 Interference Coatings

The fabrication of interference coatings by sol–gel techniques on glass is very common for many purposes, for example anti-reflective coatings (AMIRAN 1998), and interference coatings on halogen bulbs. However, most of the coatings are produced by physical techniques like sputtering or waver condensation methods (such as for eyeglass lenses). This is because the technology for mass fabrication is considered to be more difficult than for physical techniques. One of the drawbacks of sol–gel techniques is that, in general, each coated layer has to be densified and prefired before the next layer can be employed. This means that a lot of time and energy has to be used in making multilayer coatings. For that reason, a technology has been developed which takes advantage of polymerizable nanoparticles, surface modified with methacrylic or epoxy groupings. Coatings fabricated from these sols can be photopolymerized a very short time after dipping or spraying, and are then insoluble against the next coating step. After the fabrication of an appropriate set of green layers, a one-step firing process can be employed on glass, which leads to multilayered interference systems. This is schematically shown in Fig. 9 with SiO_2 and TiO_2.

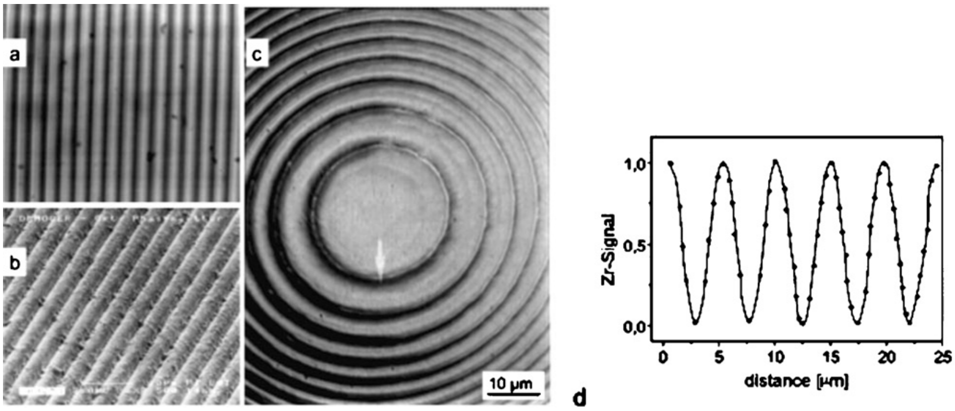


Figure 8
Microoptical component made by holographic process: (a) volume phase grating, (b) relief grating, (c) Fresnel lens, and (d) EDX line scan over a holographically fabricated diffractive grating according to (a).

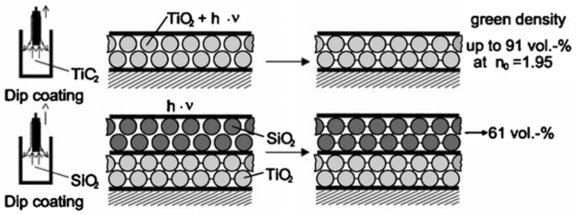


Figure 9
Scheme of the fabrication of interference layers by polymerizable nanoparticles (after Schmidt 1999).

Using this approach, a technology has been developed to produce anti-reflective layers for automotive glazings (Mennig *et al.* 1999). The technology can be employed also on plastics, as long as the photocuring process only is used. Due to the high package density of the "modified gels" (up to 90% of the theoretical density after UV-radiation), interference layers have also been obtained alternatively by using TiO_2 and SiO_2. Such coatings can be produced also on conventional hard coatings, and then show extremely high abrasion resistance. It can be shown also that the theoretical curves fit very well with the curves achieved in practice, by using the Landau–Lewich approach for calculating layer thicknesses. Taking into account the organic content, the calculations could be used for fired coatings on glass. This means that it is possible to design interference systems.

By using thicker layers, IR reflective coatings have been fabricated, which is an interesting approach for IR-shielding on plastics: for example, in the freezing industry. In this connection, it has to be mentioned that the dispersion of ITO nanoparticles, fabricated according to the processes described above, also leads to highly efficient IR-shielding coatings.

Using the nanocomposite approach, another type of anti-reflective coating was developed by the so-called "moth eye" process in nanomer surfaces combined with perfluorinated surfaces. The moth eye process leads to long-lasting anti-reflective systems only if the surface is equipped with easy-to-clean properties. Otherwise, the roughness in the nanometer range of the surface leads to dirt absorption filling out the voids and reducing the anti-reflective effect.

3. Functional and Mechanical Coatings

3.1 Hard Coatings

(*a*) *Thermally curable coatings.* The nanocomposite strategy for hard coatings is based on the fact that well-dispersed nanoparticles in hybrid matrices are considered to play a very important role in abrasion resistance. Schmidt *et al.* 1998 showed that boehmite particles, dispersed in an inorganic–organic hybrid network composed of epoxy silanes as basic components, catalyze epoxy polymerization at curing temperatures above 100 °C, leading to a highly transparent, very scratch-resistant coating. These coatings are successfully employed for the fabrication of abrasion resistant CR-39 eyeglass lenses, and have been developed for hard coatings on polycarbonate. The coating leads to very high abrasion resistance (2% haze, or 98% transparency, after 1000 cycles of taber abrader). For this reason, the coating has currently been further developed for the use of polycarbonate for automotive glazing, substituting for inorganic glasses.

(*b*) *UV-radiation curable coatings.* If UV or electron beam curing is employed (by using photocurable organic groupings like epoxides or acrylics in conjunction with appropriate photoinitiators), photocurable hardcoatings are possible. Since these coatings are produced at low temperatures (around 60 °C), the degree of inorganic network condensation is lower than in the case of temperature-cured hard coatings, leading in general to a lower abrasion resistance; but by using 10–15% of nanoparticles (by volume), the taber abrasion numbers are reduced also to about 7% after 1000 cycles, an extraordinarily high value for UV-cured hard coatings. Based on these results, a coating system for plastic lenses has been developed and an industrial process has been built up. Similar results have been obtained by electron beam curing.

By appropriate control of the drying conditions, a dry but uncured photocurable system is obtained, which has sufficient deformability to be used in deep drawing processes. By coating PMMA plates, deep drawing processes have been performed in which subsequent UV-curing leads to highly abrasion-resistant coatings on PMMA, and an industrial process has been developed.

3.2 Low Surface Free Energy Hard Coatings

Low surface free energy hard coatings are a very interesting area of application for hybrid matrix nanocomposites. As described elsewhere (Schmidt *et al.* 1991, Kasemann *et al.* 1992a, 1992b, Kasemann and Schmidt 1994), it could be shown that, based on interfacial thermodynamics, a self-alignment in layers takes place if side chain perfluorinated silanes are added, which increase in concentration at the surface as long as the concentration exceeds the critical micelle concentration. This leads to very low surface free energies of about 19 mJ m^{-2} (20–22 mJ m^{-2} with PTFE). The process is shown schematically in Fig. 10.

Based on this effect, various types of low surface free energy coatings have been developed (high- and low-temperature curing, UV curing, and two-component systems). All of these coatings form a gradient layer with an increased concentration of fluorine at the atmospheric side, as shown by ESCA or SNMS, and an adhesive layer to the substrates. Due to the homogeneous distribution of nanoparticles, these systems can also be obtained with high transparency and high abrasion resistance. The contact angles of this type of coating with water are of the order of 110°, and with hexadecane in the range of 55–60°. If micro-patterned surfaces are used, the contact angles with water increase up to 160°, and with hexadecane up to 120°.

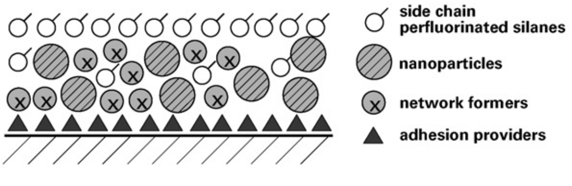


Figure 10
Scheme of a fluor gradient material.

These surfaces can be obtained either by mechanical micropatterning or by the use of self-structuring effects based on appropriate nanoparticle sizes. Based on this technology, many industrial applications have been developed. These extend through anti-adhesive coatings; mold release coatings; zero-carryover coatings in capillaries, food technology, buildings and pavements; and anti-billposting coatings—all having the advantages of saving detergents and water. The coating material can be adjusted for glass, ceramics, enamel, polymers, and metals, with average thickness ranging from 1–5 µm. A special variation of the system is the easy-to-clean application and anti-graffiti protection in public areas (house and public building walls, monuments, railways, subway, traffic signs, etc.).

Similar coatings have been developed for mold release processes in metal molds for polyurethane, polyethylene and polystyrene foams. Other applications have been developed for anti-adhesive coatings in carpet production, for easy-to-clean windows, and for ceramic sanitary ware which has been commercialized extremely successfully (Duravit Wondergliss coatings).

Another very interesting development involves the restoration of historic books and papers. Due to the content of sulfuric acid of paper (starting about 150 years ago, as a consequence of the change in production technologies), the paper slowly deteriorates, by hydrolytic decay of the cellulose fibers. This leads to a weakening of the papers, so that in serious cases the paper sheets may even become powderized. Many attempts have been made around the world to solve this problem, for example by leaching out the acid (Batelle process), or by using plastic foils as support. In the first case no strengthening is added, and in the second case a single-sheet process has been used, by dismantling the book and then binding it after fixing the foil support. For these reasons, investigations have been carried out with the aim of combining both processes, using an infiltration process with a polymerizable resin which shows a ‘‘smart’’ behavior, by knowing that it has to reinforce the cellulose fibers but at the same time prevent two pages from sticking together. The problem has been successfully solved by using the self-aligning fluorinated type of coating, in combination with acrylic groupings and a thermally activated initiator, by simply impregnating the book with a reinforcing system, drying it at temperatures of about 60–80 °C (starting temperature of the radical initiator) and then drying it. The scheme of this process is shown in Fig. 11, the mechanism of the formation of an anti-adhesive surface between two paper sheets is also depicted.

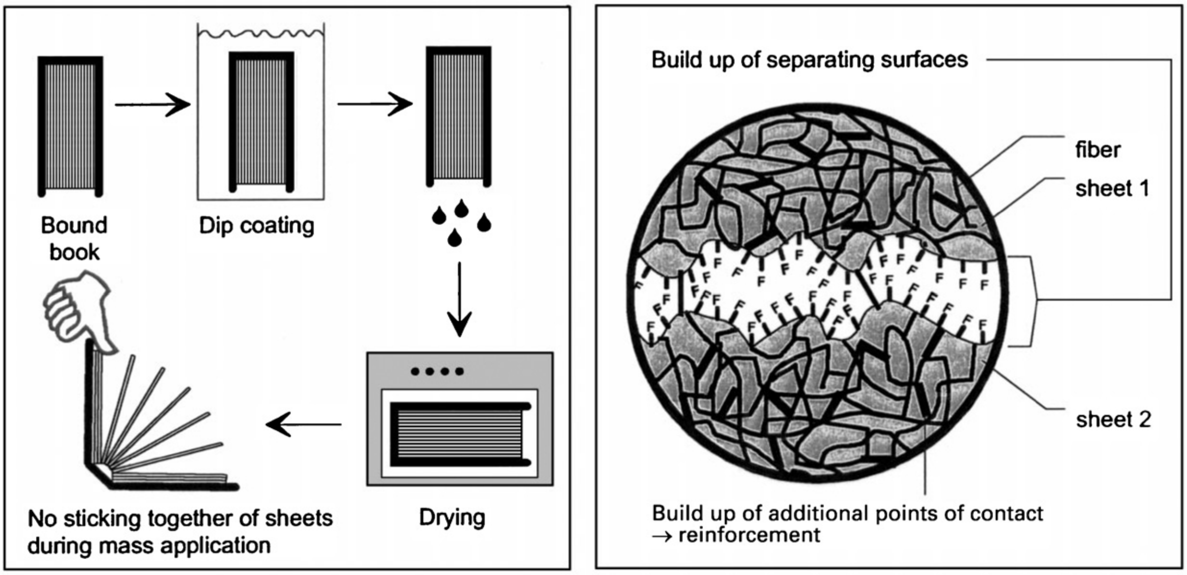

Bound book
Dip coating
Drying
No sticking together of sheets during mass application
Build up of separating surfaces
fiber
sheet 1
sheet 2
Build up of additional points of contact → reinforcement

Figure 11
Scheme of paper coating: build-up of separating surfaces and reinforcement.

3.3 Photocatalytic Coatings

TiO_2 nanoparticles show photocatalytic activities, as has been known for a long time. Suspensions of TiO_2 have been used for photocatalytic water cleaning processes. Interesting applications of TiO_2 nanoparticles have been reported by Watanabe *et al.* (1999a, 1999b), in which TiO_2 is wet coated on ceramic tiles and sanitary ware for self-cleaning and microbiocide purposes. Similar systems have been developed, based on work by Schmidt *et al.* 1995. Watanabe *et al.* 1999a showed that TiO_2 compositions have a self-cleaning effect on various substrates, especially on plastic foils, organic paints, and polycarbonate sheets.

Another area of catalytic applications is so-called non-selective catalysts, as used in oxidation catalysis. It could be shown that by the use of nanoparticulate SiO_2-containing binder materials, transition metal oxides such as those of cobalt, nickel, and others show enhanced catalytic activities compared to systems made from pellets out of the pure oxides. This type of nanocomposite has been successfully used in exhaust systems for the deodorization of kitchens.

3.4 Conductive Coatings

Electrically conducting transparent coatings are of interest for the development of antistatic coatings, heat insulating coatings, electromagnetic shielding, and displays (touch screen panels). Polymer-based wet coatings again offer the possibility of cheap, large-area coating technologies, which up until now have not been achieved by sputter techniques. It could be shown that by use of the surface modified nanoparticle approach, thicker films and lower sheet resistances could be obtained, as compared to the classical sol–gel route. The development of transparent electrical conducting coatings of ZnO:Al (AZO), SnO_2:Sb (ATO), and In_2O_3:Sn (ITO) with a defined sheet resistance, high optical transmission and good homogeneity, using the sol–gel process, led to coatings which can be obtained by dip- or spin-coating processes using stabilized sols. The sheet resistance can be adjusted by varying the layer thickness and the sintering temperature. These systems provide an adjustable sheet resistance ($\pm 2\%$) with high homogeneity of the thickness ($\pm 2\,\mathrm{nm}$). The sheet resistance can be minimized, down to 150–200 Ω on float glass at 550 °C for ATO and AZO, and on borosilicate and silica glass at 600–650 °C. For an ITO double layer, the sheet resistance is below 200 Ω, sintered at 900–1000 °C on silica glass. Substrate sizes of $1\,\mathrm{m}^2$ have already been dip coated. The temperature stability ranges up to 800 °C for ATO coatings, and up to 900 °C for ITO coatings. In the visible range, a high transmission could be derived (Fig. 12).

3.5 Metal Colloid-containing Coatings

Metal colloids have been used for two different purposes. As known from gold ruby glass, metal colloids show high coefficients of extinction due to plasmon frequencies. The coefficients of extinction are up to

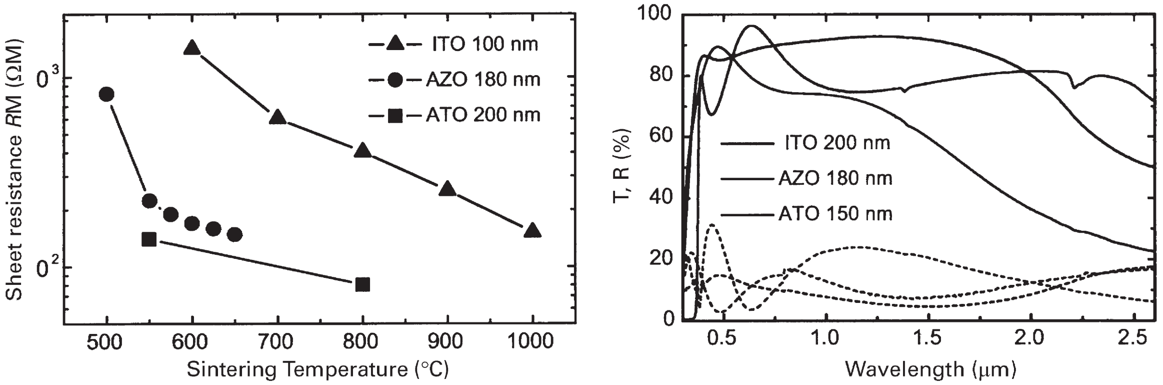


Figure 12
(left) Sheet resistance vs. sintering temperature of dip-coated SnO_2: Sb (ATO, 10 layers), In_2O_3 (ITO, 2 layers) and ZnO: Al (AZO, 10 layers) on silica glass; (right) transmission (—) and reflectance spectra (- -) of dip-coated ATO, AZO and ITO coatings.

$10^6 \, l \, mol^{-1}$ which even exceeds those of highly absorptive organic dyes. If these colloids are developed within a glass matrix by nucleation and growth, intensive colors can be obtained within thin sol–gel coatings. Based on this technology, an industrial process has been developed for the coloration of silicate eyeglass lenses, using different metals for different colors (gold colloids for red, silver colloids for yellow, copper colloids for green, and palladium colloids for brown). In the industrial process (Mennig *et al.* 1998b), a lead silicate glass matrix is doped with metal colloid precursors in a complexed form, and by a spin coating process a 200 nm thick coating is applied. After firing these coatings under reductive conditions, nucleation and growth take place, followed by the formation of the glass-like colored coatings. Mechanical properties of these coatings are similar to those of the substrate glass. Another interesting application of these particles is in non-linear optical materials based on gold particles (Mennig *et al.* 1994a). The fabrication of nanoparticle-containing coatings can be carried out by doping of the sol–gel precursors with gold complexes, and gold colloids can be grown in a glass matrix as well as in hybrid matrices. Other developments for materials with high χ^3-nonlinear optical effects have been obtained by growing gold nanoparticles in hybrid matrices. For this process, precursor complexes from aminosilanes with $AuCl_3$ were synthesized. After dispersing these complexes in TEOS and methacryloxysilane-containing systems, coating liquids were fabricated by hydrolysis and condensation. After coating and drying, photopolymerization of the methacrylic groupings was carried out. During this period a nucleation and growth process of gold particles took place, and gold particles of several nm in diameter were produced (Mennig *et al.* 1994b). The χ^3 coefficient of $\chi^3/\alpha \approx 10$ (10^{-14} esu m) could be obtained.

3.6 Coatings for Corrosion Protection

The possibility of the molecular design of polymer-based composites on the nanometer scale allows the adhesive forces to the substrate to be optimized. Various systems have been developed in order to stop corrosion processes on metallic substrates. The corrosion process itself occurs in the presence of an electrochemical couple, i.e., in the presence of metals that have a high tendency for the exchange of electrons. This setup is equivalent to an electrochemical cell, with electron donors and acceptors. Due to the change of the ionic state at the metal surface, chemical reactions take place (passivation). Metal oxides are formed, which lower the metallic properties of the components (strength, toughness, color, etc.). The transport of electrons takes place effectively in the presence of a polar liquid, e.g., water. Therefore, an effective corrosion protection system will prevent contact of water with the metal surface and so prevent electron transfer and the formation of corrosion products. Polymer-based nanocomposites have been developed which provide excellent adhesion forces, which can be tailored to various metals.

Additionally, anti-adhesive "easy-to-clean" properties can be built into these coatings. Many other systems have been developed, e.g., as tarnish protection on silver, brass, and magnesium. The nanocomposite works without primer and has good adhesion to metals, offering high mechanical resistance at low coating thickness (approx. 5–10 μm). A special development is based on surface-modified SiO_2 nanoparticles, as described in Sect 3.2. These films can be densified at 500–600 °C. Highly scratch-, tarnish-, and corrosion-resistant coatings for stainless steel have been prepared. The coatings are 3–5 μm thick and very flexible. Another coating composition has led to anti-fingerprint coatings on stainless steel.

Further advantages of this system lie in the application as an electrical isolation layer for high temperature applications on stainless steel, mild steel, copper, brass, and aluminum alloys. Compared to hardening of steel chromium plating, the glass coating shows an increased wear resistance. The decorative appearance of glossy steel surfaces is obtained up to 600 °C.

4. Hybrid Nanoparticles

4.1 Processing Aids

As described before, the surface modification of hybrid particles is one of the key issues in processing these particles to obtain desired materials and desired functions. In many cases, the particle consists of an inorganic core and an organic shell. These particles can also be defined as hybrid nanoparticles. This definition is used if the hybrid particle plays an important role in processing or in the final properties. In the field of ceramics, sintering temperature and mechanical performance play an important role. Many investigations have been devoted to reducing the sintering temperatures by reducing the particle size. However, in most cases the package density becomes a serious problem. For this reason, it was investigated how to use hybrid nanoparticles as ceramic precursors. As shown by Nass *et al.* (1995), with titanium nitride as an example, nanotitanium nitride coated with guanedinocarboxylic acid leads to package densities in colloidal shaping processes up to 60% by volume. These relatively densely packed matrices could then be densified to titanium nitride compacts or coatings at temperatures from 1200–1350 °C, showing that by hybrid nanoparticle processing, substantial reductions of sintering temperatures could be obtained; and it could be shown also that grain growth could be "separated" from the densification step. The grain growth takes place mainly above 1350 °C.

In the case of silicon carbide, it could be shown that nanometer-size carbon particles coated with carboxylic acids could be linked to silicon carbide surfaces containing amino groups, in such a way that wet processing did not lead to bleeding of the carbon black during slip casting. By using an appropriate concentration of carbon black, suitable for the stoichiometric reduction of the SiO_2 always present on silicon carbide particle surfaces, a slip casting processing route was developed which leads to near net-shaped large silicon carbide parts (Nass *et al.* 1994). Silicon carbide igniters based on this technology shows superior properties to those previously available (500 000 ignition cycles without rupture).

4.2 Medical Applications

Nanoparticles are interesting for medical applications for many reasons. One is that for *in vivo* applications they can be transported in the blood stream without any problems, can even pass important barriers and are not filtered out very easily. Moreover, with appropriate surfaces they may penetrate even cell membranes. Based on these considerations, amino-grouping-modified superparamagnetic iron oxide nanoparticles have been fabricated, and investigations have been carried out for transporting them into tumor cells (Lesniak *et al.* 1998). It has been shown that billions of nanoparticles penetrate tumor cells. By employing an alternative magnetic field, heat could be introduced into the cells and they could be selectively destroyed by heat. It was also shown (H. Jordan, personal communication 1998) that if the tumor cells divide, both daughter cells share the load of iron oxides.

Another technology is the use of SiO_2 nanoparticles, the surface charge of which is tailored by amino grouping modification in such a way that genes can be concentrated on top of the nanoparticles, leading to gene–plasmid clusters of about 80 nm in diameter. This is a substantial reduction from the original gene plasmids with a diameter of 200–250 nm. At this size, the nanoparticles could be used for the transfection of gene plasmids into cells with a very high transfection rate (Kneuer *et al.* 1999). The advantage of this technology is the fact that the SiO_2 is completely nonpoisonous compared to polyethylene imines as gene carriers. An example of this coupling is the electrostatical bonding of glycoproteins (e.g., avidin) on top of a negatively charged SiO_2-nanoparticle. Avidin provides four sites for the chemical bonding of other groupings (like biotin/WGA-complexes) or DNA-fragments or genes (like WGA: wheat germ agglutinin). The aim of this approach is the destruction of cancer cells by pharmaceuticals. The coupling of biomolecules can be proved by electron microscope investigations.

Hydroxylapatite has been fabricated from nanoparticulate calcium phosphate precipitates and, using sol–gel binders, these nanoprecipitates have been used for fabricating thin films on titanium implants in thicknesses of about 5 nm. This thickness has great advantages compared to plasma- or flame-sprayed thick films, because it adheres very well on the surface and no cracks are formed. First results in using these implants show that a very fast incorporation into the tissue takes place (Nonninger *et al.* 2001).

5. Conclusion

Nanoparticles are of great interest for many applications, either for use by themselves or by being processed to parts such as in ceramics, or by being used in the form of nanocomposites in various matrices. Because nanoparticles of many different compositions can be used, a wide range of properties can be developed for materials. Chemical fabrication routes for nanoparticles are very interesting, since they are mainly based on precipitation techniques which can be performed quite cheaply. One of the key issues is the surface modification of nanoparticles, since that is the crucial point in controlling aggregation. The application of these principles to the development of materials will be a very important and interesting prospect for the future.

Bibliography

AMIRAN 1998. *Technical information.* Schott Glaswerke

Hou L, Mennig M, Schmidt H 1994 Photochromic organic-inorganic composite materials prepared by sol–gel processing: properties and potentials. *Proc. SPIE-Int. Soc. Opt. Eng.* **2255**, 26–37

Judeinstein P, Oliveira P W, Krug H, Schmidt H 1997 Photochromic organic–inorganic nanocomposites as holographic storage media. *Adv. Mater. Opt. Electron.* **7** (3), 123–33

Kasemann R, Brueck S, et al. 1992a Fluorine modification of ORMOCER (organically modified ceramics)-coating materials. *Prog. Res. Dev. Process. Prod. Sol-Gels.* **5**, 353–61

Kasemann R, Brueck S et al. 1992b Schmutzabweisende, transparente Beschichtungen auf Basis fluormodifzierter anorganisch–organischer Nanokomposite. *Glastechnische Tagung*. DGG

Kasemann R, Schmidt H 1994 Self-assembling fluorine-containing nanocomposite coatings by sol–gel techniques. In: *Int. Conf. Fluorine in Coatings*

Kneuer C, Sameti M, Haltner E, Schiestel T, Schirra H, Schmidt H, Lehr C M 1999 New materials for gene delivery: surface modified silica nanoparticles can efficiently bind, protect and transfect genes in-vitro. In: *CRS Annu. Meeting*.

Lesniak C, Schiestel T, Schmidt H, Jordan H 1998 *WO 9858673 A1*

Lukosz W, Tiefenthaler K 1983 Embossing technique for fabricating integrated optical components in hard inorganic waveguiding materials. *Opt. Lett.* **8** (10), 37–9

Mennig M, Jonschker G, Schmidt H 1992 Sol–gel derived thick SiO_2 coatings and their thermomechanical and optical properties. *Proc. SPIE.* **1758**, 125–34

Mennig M, Munro B, Koch T, Zapp P, Heusing S, Schmidt H 1998a Thick film polymerisable electrolytes for electrochromic glazing. In: *Proc. 2nd ICCG*

Mennig M, Oliveira P W, Frantzen A, Schmidt H 1999 Multilayer NIR reflective coatings on transparent plastic substrates from photopolymerizable nanoparticulate sols. *Thin Solid Films* **351**, 225–9

Mennig M, Schelle C, Duran A, Damborinea J J, Guglielmi M, Brusatin G 1998b Investigation of glass-like sol–gel coatings for corrosion protection of stainless steel against liquid and gaseous attack. *J. Sol–gel Sci. Technol.* **13**, 717–22

Mennig M, Schmitt M, Becker U, Jung G, Schmidt H 1994a Gold colloids in sol–gel derived SiO_2 coatings on glass and their linear and nonlinear optical properties. *Proc. SPIE* **2288**, 130–9

Mennig M, Schmitt M, Becker U, Ulrich G, Jung G, Schmidt H 1994b SiO_2 coatings on glass containing copper colloids using the sol–gel-technique. *Proc. SPIE* **2288**, 120–9

Munro B, Conrad P, Krämer S, Schmidt H, Zapp P 1998 Development of electrochromic cells by the sol–gel process. *Solar Energy Mater. Solar Cells* **54**, 131

Nass R, Albayrak S, Aslan M, Schmidt H 1995 Processing and sintering of nanosized TiN. *Proc. CIMTEC.* **11**, 47–54

Nass R, Aslan M, Nonninger R, Schmidt H, Matje P 1994 New processing techniques for the production of pressureless sintered SiC parts. *Proc. Int. Conf. Ceram. Proc. Sci. Tech. ACerS.* **51**, 433–7

Nonninger R, Sauer M, Schmidt H 2001 Bioactive hydroxyapatite coating for titanium implants. In press

Oliveira P W, Krug H, Müller P, Schmidt H 1996 Fabrication of GRIN-materials by photopolymerization of diffusion-controlled organic–inorganic nanocomposite materials. *MRS Symp. Proc.* **435**, 553–8

Schmidt H 1996 Relevance of sol–gel methods for synthesis of fine particles. *Kona Powder Particle* **14**, 92–103

Schmidt H 1999 Sol–gel derived nanoparticles as inorganic phases in polymer-type matrices. In: *Proc. Int. Conf. Pol. Sci.* In print

Schmidt H, Geiter E, Mennig M, Krug H, Becker C, Winkler R P 1998 The sol–gel process for nano-technologies: new nanocomposites with interesting optical and mechanical properties. *J. Sol–gel Sci. Techol.* **13**, 397–404

Schmidt H, Kasemann R et al. 1991 German Patent 4 118 184 A1

Schmidt H, Langenfeld S, Naß R 1997 A new corrosion protection coating system for pressure-cast aluminum automotive parts. *Mater. Des.* **18** (4/6), 309–13

Schmidt H, Mennig M et al. 1995 German Patent 43 38 360 A1, 1996 European Patent 072 9 442 B1

Schmidt H, Nonninger R 1998 Chemical routes to nanoparticles: synthesis, processing and application. In: *Proc. Fine, Ultrafine and Nano Powders*

Tohge N, Shinmou K, et al. 1994 Photolysis of organically modified gel films and its application to the fine-patterning of oxide thin films. *Proc. SPIE* **2288**, 589–98

Watanabe T et al. 1999a Japanese Patent 11076833 A

Watanabe T et al. 1999b Japanese Patent 11153701 A

H. Schmidt and M. Stadtwald-Klenke
Universität des Saarlandes, Saarbrucken, Germany

Polymer–Ceramic Nanocomposites: Catalysts

Polymers can be incorporated into ceramics at a molecular level using the sol–gel process. In this, a ceramic phase is generated by an inorganic polymerization reaction of forming a metal oxo (M-O-M) polymer using organometallic precursors. The reaction generally includes two steps: (i) the hydrolysis of molecular precursors such as metal alkoxides $M(OR)_Z$, where M is a metal or transition metal such as Si, Ti, Zr, Al, Sn, Ce, etc., R is alkyl, phenyl, or hydrogen, and Z is the valence or the oxidation state of the metal; and (ii) the condensation of the hydroxyl groups and the residual alkoxyl groups to form a three-dimensional network. Once the hydrolysis reaction has been initiated, both hydrolysis and condensation reactions occur simultaneously. Depending on the relative reaction rates, the structure and morphology of the ceramic phase will be different. Many chemical species such as acids, bases, and some neutral salts have a large impact on the reaction mechanisms and the relative reaction rates, and can thus be used as catalysts in the process of making polymer–ceramic nanocomposites through the sol–gel process.

1. Catalytic Mechanisms

The sol–gel process and the reaction mechanisms have been extensively studied since the early 1980s (Brinker and Scherer 1990). However, a brief discussion on the reaction mechanism is helpful to explain the catalytic effects.

In the absence of catalysts, hydrolysis and condensation both occur by nucleophilic substitution (S_n) mechanisms involving nucleophilic addition (A_n) followed by proton transfer from the attacked molecule to an alkoxide or hydroxo-ligand within the transition state. This is then followed by the removal of the protonated species as either alcohol (alcoxolation) or water (oxolation) (Scheme 1).

When the coordination of the metallic center is not fully saturated, condensation can occur by olation based on a nucleophilic addition mechanism:

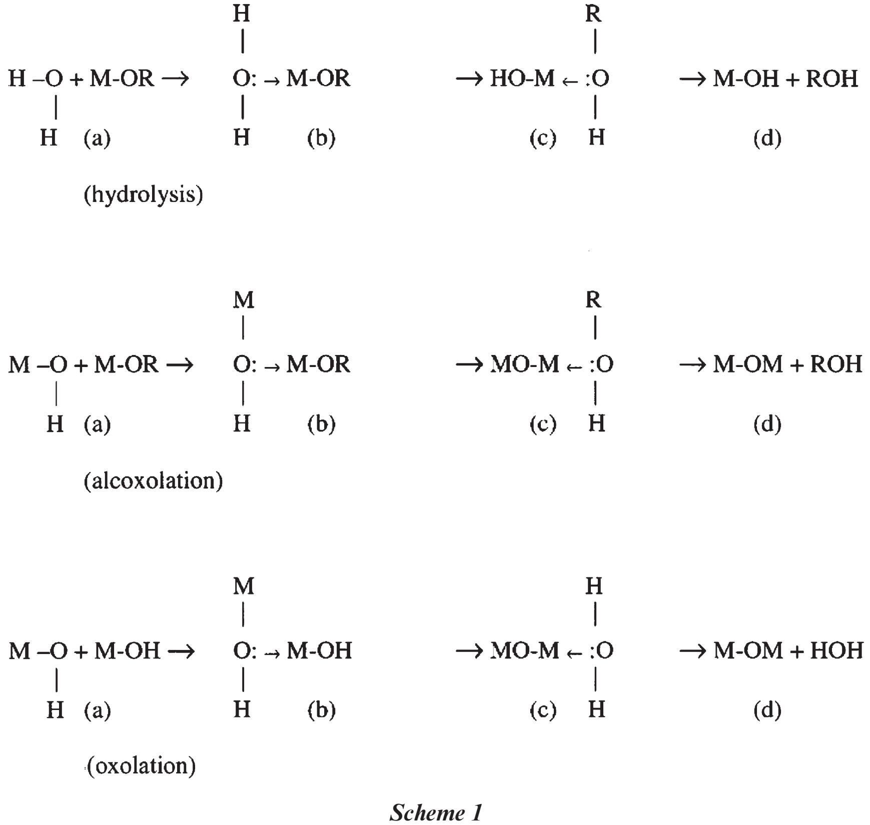
$$\mathrm{MOX} + \mathrm{MOX} \rightarrow \mathrm{M} - (\mathrm{OX})_2 - \mathrm{M}$$

$$\mathrm{X}:\mathrm{H} \quad \text{or} \quad \mathrm{R}$$

(olation)

The thermodynamics of hydrolysis, alcoxolation, and oxolation are governed by the strength of the entering nucleophile, the electrophilicity of the metal, and the partial charge and the stability of the leaving group. The kinetics of these reactions are governed by the extent of coordination undersaturation of the metal, and the transfer ability of the proton. Greater acidities of the protons reduce the associated activation barriers and enhance the kinetics. The thermodynamics of the olation depends on the strength of the entering nucleophile and the electrophilicity of the metal, and the kinetics of olation is systematically fast. This is because of the undersaturation, and the fact that no proton transfer occurs in the transition state.

Any of these reactions can take place in the sol–gel process. Depending on the relative contribution of each of these reactions, the structure and the morphology of the resulting ceramic phase are different.



Scheme 1

The nature of the metal atom, the alkyl groups, and the molecular complexity are the internal factors. Catalysts as well as the water ratio, concentration, solvent, and temperature are external factors that can also be used to control the reactivity and thus tailor the structure of the composite.

The chemical reactivity of metal alkoxides mainly depends on the electrophilic character of the metal atom (electronegativity), and its ability to increase its coordination number. The reactivity of the commonly used tetravalent metal alkoxides has the sequence

$$\mathrm{Si(OR)_4} <<< \mathrm{Sn(OR)_4} < \mathrm{Ti(OR)_4} < \mathrm{Zr(OR)_4} < \mathrm{Ce(OR)_4}$$

The reactions of silicon alkoxides are much slower than those of nonsilicon alkoxides. Catalysts are normally needed in silicon–alkoxide systems. In non-silicon systems, inhibitors or chelating ligands are sometimes used instead, to control or slow down the sol–gel reactions of these alkoxide precursors (Livage *et al.* 1988).

During hydrolysis, the acids protonate negatively charged alkoxide groups. Electron density is then withdrawn from the metal atom, making it more electrophilic and thus more susceptible to attack by water. It enhances the reaction kinetics by producing good leaving groups. Under basic conditions it is likely that water is first dissociated to produce nucleophilic hydroxyl anions. The hydroxyl anion then attacks the metal atom. Although both acid and base conditions promote the hydrolysis, acids are more effective.

During condensation, it is generally believed that the acid also protonates the MOH species which are more electrophilic and more susceptible to nucleophilic attack. Basic conditions on the other hand produce nucleophiles via deprotonation of the hydroxo ligand, which systematically enhances the condensation. The condensation pathway is influenced by the relative ease of protonation of different alkoxide ligands. The ease of protonation is related to the electron-providing power of the ligands, which decreases in the order alkoxy (-OR), hydroxo (-OH), and oxo (-O-). Under acidic condensation, easily protonated ligands such as OR and OH are favorable, thus the condensation reaction occurs preferentially among them, resulting in more-extended, less-branched ceramic phases. Under basic conditions, the more acidic ligands, such as oxo, are favorable, therefore the final ceramic phase is more compact and highly branched. Many different reaction mechanisms have been proposed for these catalytic conditions (Brinker and Scherer 1990). In addition to acids and bases, some salts, especially F^- salts, are very effective catalysts. Their catalytic effect suggests that F^- may have the same role as OH^-.

2. Applications

Many acids, bases, and other species have been used to catalyze the sol–gel process for preparing polymer–ceramic nanocomposites. Substances with high volatility, such as ammonia and diethylamine, are normally used to permit their easy removal from the resulting systems. Acids, bases, and some neutral salts have been used to promote the generation of silica particles in an elastomeric network using metal–alkoxide precursors to improve its mechanical properties (Mark 1989). The morphology of the resulting silica-filled polydimethylsiloxane (PDMS) network strongly depends on the catalysts used. It has been found that silica generated in PDMS elastomer networks has a good particle definition with a smooth surface when the catalyst is a base, whereas an acid catalyst gives poorly defined fuzzy particles. The latter may be an advantage with regard to good interfacial bonding. The mechanical properties of the resulting composites are correspondingly different.

Another type of polymer–ceramic hybrid has been prepared from simultaneous formation of silica networks and PDMS networks. The resulting hybrids can be very rubbery when a polymer is the continuous phase, or ceramic-like with polymer as a toughing agent. The acid content is very crucial in determining the structure of the final products when oligomeric PDMS is incorporated into inorganic glass made by the sol–gel process of tetraethoxysilane (TEOS) (Huang *et al.* 1987). As the acid content increases, PDMS is better dispersed into the system. Higher polymer content can be obtained in making transparent homogeneous hybrid materials.

When several different metal–alkoxide precursors are used in the process, catalysts can also be used to match or control the reactivities among the precursors, and thus to tailor the structure and properties of the composites. One example is that the hydrolysis rate of a methyl-substituted silicon alkoxide is much faster than that of a pure silicon alkoxide in an acidic medium, but the rate is reversed in a basic one. The reaction medium thus can be used to match the reactivities of alkoxysilane and methyl-substituted alkoxides (Schmidt 1984). Different pH values can also be used to tailor the compositions of the resulting composites when different metal alkoxides are used (Dire *et al.* 1992).

Polymeric acids can also be employed, and the choice of its molecular weight can be used to control the viscosity of the reacting system (Brennan and Wilkes 1991).

The sol–gel process makes it possible to make polymer–ceramic composites at the molecular level. The use of acids, bases, and many other chemical species can be used to control reaction rates and reaction pathways. In this way, the structure and morphology of the resulting polymer–ceramic nanocomposites can be tailored.

Bibliography

Brennan A B, Wilkes G L 1991 Structure–property behavior of sol–gel derived hybrid materials: effect of a polymeric acid catalyst. *Polymer* **32**, 733–9

Brinker C J, Scherer G W 1990 *Sol–Gel Science, The Physics and Chemistry of Sol–Gel Processing*. Academic Press, San Diego, CA

Dire S, Babonneau E, Sanchez C, Livage J 1992 Sol–gel synthesis of siloxane–oxide hybrid coatings. *J. Mater. Chem.* **2** (2), 239–44

Huang H, Orler B, Wilkes G L 1987 Structure–property behavior of new hybrid materials incorporating oligomeric species into sol–gel glasses. 3. Effect of acid content, tetraethoxysilane content, and molecular weight of polydimethylsiloxane. *Macromolecules* **20**, 1322–30

Livage J, Henry M, Sanchez C 1988 Sol–gel chemistry of transition metal oxides. *Prog. Solid State Chem.* **18**, 259–341

Mark J E 1989 Generate reinforcing particles in place. *CHEMTECH* **19**, 230–3

Schmidt H 1984 Organically modified silicates by the sol–gel process. In: Brinker C J, Clark D E, Ulrich D R (eds.) *Better Ceramics Through Chemistry*. North-Holland, New York, pp. 327–35

S. Wang
DuPont Dow Elastomers, Elkton, Maryland, USA

Polymer–Ceramic Nanocomposites: Ceramic Phases

The sol–gel process has been extensively studied and employed as the most important route in forming new hybrid inorganic–organic materials (Mark 1996, Mark *et al.* 1995, Sharp 1998, Sanchez and Ribert 1994, Novak 1993, Chujo and Saegusa 1992, Schmidt 1994, Schaefer and Mark 1990). This is because this process can introduce ceramic phases into composites through chemistry. Very mild reaction conditions and low reaction temperatures are particularly useful for incorporating inorganic moieties into organic materials or organic materials into inorganic matrices. Thus, the organic polymers involved can survive during the sol–gel process that eliminates sintering temperatures used in the preparation of typical inorganic ceramics ($>1200\,^{\circ}$C).

The sol–gel process begins with alcoholic solutions of metal or semi-metal alkoxide precursors $M(OR)_n$ and water, where M is a network-forming element such as Si, Ti, Zr, Al, B, etc. and R is typically an alkyl group. Metal alkoxides are compounds in which elements are bonded to hydrocarbon groups through oxygen.

Hydrolysis and condensation of metal alkoxides can lead to the formation of inorganic networks in the presence of an acidic or basic medium. The natures of the catalysts and pH in the sol–gel reaction have very significant effects on the hydrolysis and condensation reactions and the morphology of the inorganic network formed. Such a sol–gel process can lead essentially to an alloy-like material if molecular dispersion is obtained or, in most cases, more of a nanophase morphological structure. It is known that physical and optical properties of inorganic–organic composites are enhanced by the reduction of the domain size. As the domain size is reduced, the interfacial surface area increases. Substantial improvement in performance can be achieved with domains on the nanometer scale.

1. Precursors for the Formation of Ceramic Phases

Some metal alkoxides used as precursors for the sol–gel process to form ceramic phases and their molecular association degrees are listed in Table 1. They may be considered either as derivatives of alcohols or as derivatives of metal hydroxides or inorganic acids (Bradley *et al.* 1978). The molecular association or oligomerization of metal alkoxides is a very important parameter that can influence many of the physical and chemical properties of these compounds and affect the sol–gel process. The degree of oligomerization can directly induce differences in processing and in final morphologies of hybrid inorganic–organic composites, such as inorganic phase domain sizes and phase separations. The degree of the molecular association depends on the nature of the central metal elements and on the nature of the ligands. When the same alkoxy ligand is used, for example, $(OEt)_4$, it has been found that the degree of oligomerization is a function of the metal size in the order Th > Hf > Zr > Ti > Si. It has also been found that the degree of oligomerization is a function of the length and branching of the alkoxy ligand (Guglielmi and Carturan 1988).

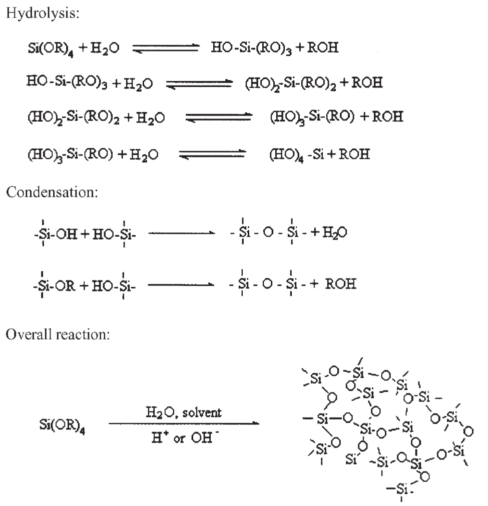
2. Metal Alkoxides Commonly Used for Hybrid Inorganic–Organic Nanocomposites

2.1 Silicon Alkoxides

Silica is one of the most important ceramic phases, and has been widely studied in the sol–gel process to create inorganic–organic nanocomposites. The tetraalkyl orthosilicates most commonly used in the sol–gel process are either tetramethoxysilane (TMOS) or tetraethoxysilane (TEOS). The general reaction steps are given in Scheme 1. Besides the nature of the catalysts used, pH has a dramatic effect on the reaction rates and the morphology of the resulting silica network. During hydrolysis and condensation of tetraalkyl orthosilicates under basic conditions, the rate of condensation is faster relative to the hydrolysis rate, resulting in more of a dense cluster growth leading to dense particulate-like structures. In contrast, if the

Table 1
Metal alkoxides used for forming ceramic phases and their molecular association.

Metal alkoxide		Molecular association
Group I	$LiOBu^t$	hexameric
Group II	Be primary alkoxides	infinite polymers
	tertiary alkoxides	dimeric,trimeric
Group III	B lower alkoxides	monomeric
	$Al(OEt)_3$	tetra- to hexameric when fresh polymeric on aging
	$Al(OPr^i)_3$	dimeric in vapor phase trimeric in solution when fresh tetrameric on aging
	$Al(OBu^t)_3$	dimeric
	$Ga(OEt)_3$	tetrameric
	$Ga(OPr^i)_3$	dimeric
Group IV	Si, Ge alkoxides	monomeric irrespective of the chain length and branching of alkoxide group
	$Sn(OMe)_4$	highly associated
	Sn *n*-alkoxides	tetrameric
	$Sn(OPr^i)_4$	trimeric
	$Sn(OR^t)_4$	monomeric
Group V	As alkoxides	monomeric
	Sb alkoxides	monomeric
	Bi alkoxides	monomeric
Group VI	S alkoxides	monomeric
	Se alkoxides	monomeric
	Tl alkoxides	monomeric
Transition elements		
	Co^{II}, Ni^{II}, Cu^{II} alkoxides	polymeric
	Mn^{II}, Cr^{II} alkoxides	polymeric
	$Cr(OBu^t)_3$	dimeric
	$Cr(OBu^t)_4$	monomeric
	$V(OMe)_4$	trimeric
	$V(OEt)_4$	dimeric
	$V(OPr^n)_4$	1.38
	$V(OBu^n)_4$	1.31
	$V(OAm^n)_4$	1.27
	$V(OR^s)_4$	monomeric
	$V(OR^t)_4$	monomeric
	Fe^{III} *n*-alkoxides	trimeric
	$Fe(OPr^i)_3$	dimeric
	$Fe(OBu^t)_3$	dimeric
	$Ti(OMe)_4$	tetrameric
	$Ti(OEt)_4$	2.4 depending on environment and concentration
	$Ti(OPr^n)_4$	trimeric
	$Ti(OBu^n)_4$	trimeric
	$Ti(OPr^i)_4$	1.4
	$Zr(OMe)_4$	polymeric
	$Zr(OEt)_4$	3.6
	$Zr(OPr^i)_4$	trimeric
	$Zr(OAm^s)_4$	dimeric
	$Zr(OR^s)_4$	monomeric
	$Zr(OR^t)_4$	monomeric
	Ce^{IV} *n*-alkoxides	tetrameric
	Th^{IV} *n*-alkoxides	hexameric
	Nb^V *n*-alkoxides	dimeric
	Ta^V *n*-alkoxides	dimeric


Hydrolysis:
Si(OR)4 + H2O ⇌ HO-Si-(RO)3 + ROH
HO-Si-(RO)3 + H2O ⇌ (HO)2-Si-(RO)2 + ROH
(HO)2-Si-(RO)2 + H2O ⇌ (HO)3-Si-(RO) + ROH
(HO)3-Si-(RO) + H2O ⇌ (HO)4 -Si + ROH
Condensation:
-Si-OH + HO-Si- → -Si-O-Si- + H2O
-Si-OR + HO-Si- → -Si-O-Si- + ROH
Overall reaction:
Si(OR)4
H2O, solvent
H+ or OH-

Scheme 1

hydrolysis and condensation reactions are carried out in the presence of acid, it is believed that the hydrolysis rate is higher than the condensation rate. As a result, more linear or polymer-like chain structures are formed (Brinker and Scherrer 1990) (see *Polymer–Ceramic Nanocomposites: Catalysts**).

2.2 Aluminum Alkoxides

The fundamental process that transforms aluminum alkoxides into a porous transparent alumina was developed by Yoldas (1975). An aluminum alkoxide, $Al(OR)_3$, normally $Al(OBu^s)_3$ or $Al(OPr^i)_3$, is first hydrolyzed in a large excess of water at 80–100 °C, resulting in precipitation of fibrillar boehmite, followed by peptization with a mineral acid to yield a clear sol. Cold-water hydrolysis leads to an amorphous precipitate that converts to bayerite upon aging through a dissolution–recrystallization process. Gelation is generally obtained by concentration of the sol via boiling or evaporation. The hydrolysis and condensation of aluminum alkoxides are complicated in the Yoldas process, but the overall reaction is shown simply by

$$2Al(OR)_3 + 3H_2O \rightarrow Al_2O_3 + 6ROH \qquad (1)$$

Acids and bases can influence the above hydrolysis and condensation reactions. Acids protonate OR or OH ligands creating good leaving groups and eliminating the requirement for proton transfer in the intermediate, while bases deprotonate water or OH ligands creating strong nucleophiles.

2.3 Titanium Alkoxides

In general, the hydrolysis and condensation reactions of titanates are complicated, but the overall reaction scheme is

$$Ti(OR)_4 + 2H_2O \rightarrow TiO_2 + 4ROH \qquad (2)$$

The reaction rate of normal titanium alkoxides generally decreases with increasing alkyl chain length. The reactivity of isomeric alkyl derivatives is tertiary > secondary > normal.

2.4 Zirconium Alkoxides

The process to prepare ZrO_2 from $Zr(OR)_4$ consists of at least two reaction steps:

$$Zr(OR)_4 + H_2O \rightarrow ZrO(OR)_2 + 2ROH \qquad (3)$$

$$2ZrO(OR)_2 \rightarrow ZrO_2 + Zr(OR)_4 \qquad (4)$$

The nature of the alkyl groups can significantly influence particle sizes and phase morphologies. ZrO_2 made from zirconium *n*-butoxide annealed below 300 °C is amorphous. Its local order is similar to that of the tetragonal crystallite phase. The amorphous phase can crystallize into the tetragonal and monoclinic zirconia nanophases at 300–350 °C, with the monoclinic crystals having a smaller size. Annealing zirconia at temperatures higher than 350 °C transforms the tetragonal phase into the monoclinic phase.

3. Organofunctional Metal Alkoxides

Organic functional and inorganic groups can be linked by stable chemical bonds, thus producing organofunctional metal alkoxide derivatives. There is a wide variety of ways to make organofunctional alkoxide derivatives (Schmidt 1994, Schubert *et al.* 1995). A general molecular formula, $(RO)_n$–M–X–A, can be written, where A represents an organic functional group and X is a chemically stable spacer linking A and the metal alkoxide moiety $M(OR)_n$. For example, alkoxide silanes, in which A is a reactive vinyl group, are extremely interesting for the synthesis of hybrid inorganic–organic composites by polymerizing the vinyl groups. In this approach, an organic network is formed in the matrix of the inorganic network. If A is a nonlinear optic chromophore group, inorganic– organic composites to be formed will have nonlinear optical properties. Theoretically, A can be any organic functional group. Therefore, there can be an unlimited number of organofunctional metal alkoxides.

4. Properties and Applications of Some Ceramics

Table 2 lists some physical properties of the ceramics which can be prepared by the sol–gel process along with their applications. There are many studies showing that incorporating silica into hybrid inorganic–organic composites can increase mechanical

Table 2
Some ceramic phases created in inorganic–organic nanocomposites and their properties and applications.

Ceramics	Properties	Applications
SiO_2 (from silicates)	Optical, mechanical, dielectric constant	Optical elements wave guides integrated optics
	Surface treatment	Reinforcement Surface treatment Catalyst support Insulation
TiO_2 (from titanates)	High index, heat resistance, dielectric constant, piezoelectric	Capacitors light valves
Al_2O_3 (from aluminates)	Mechanical, thermal	High modulus High thermal conductivity
ZrO_2 (from zirconates)	Mechanical, heat resistance	High modulus High temperature
GeO_2 (from germanates)	High index, magnetic	Optical fiber
P_2O_5 (from phosphates)	Low temperature	Sensors
B_2O_3 (from borates)	Optical, mechanical, chemically durable	Optics Glassware
V_2O_5 (from vanadates)	Conductive	Semiconductor conductors varistors switching devices
Nb_2O_5 (from niobates)	Electro-optic, piezoelectric	Light valves modulator hydrophone
WO_3 (from tungstates)	Electrochromic	Displays

properties such as modulus and ultimate tensile strength, and improve interfacial bonding between the organic and inorganic phases. TiO_2 and GeO_2 can be used to produce high-refractive index optical films. TiO_2 can also increase the thermal stability of composites—for example, titania prepared from titanates can overcome a thermal degradation problem caused by the silanol groups on the surface of the silica at high temperatures. Al_2O_3-containing composites can have higher hardness and scratch resistance, while zirconia-based composites are light and able to withstand heat, corrosion, and wearing. Many of these zirconia-based composites have been used for coating turbine blades and for making engine blocks, body valves, cylinder liners, pistons, and bearings in automobile internal combustion engines. This is because composites based on zirconia have a high fracture toughness produced by martensitic transformation of the zirconia tetragonal phase into the monoclinic phase. The mechanical properties of these composites depend strongly on the size of the zirconia crystallite (Komarneni *et al.* 1997).

5. *Nanocomposites Containing Various Ceramic Phases*

In the case of the inorganic–organic nanocomposites, the goal is to carry out reactions of the sol–gel type in the presence of organic molecules which are generally polymers and have functional groups to improve their interfaces with the glass phase. There is a wide variety of ways to generate such a glass phase in polymers. In general, there are two classes of hybrid inorganic–organic nanocomposites based on inorganic and organic interactions (Sanchez and Ribert 1994). In class I, organic molecules, oligomers, or low-molecular weight polymers are simply embedded in inorganic matrices. Conversely, inorganic phases can be generated *in situ* in organic matrices. In this class, both components have relatively weak van de Waals, hydrogen, or ionic interactions. In class II, inorganic, organic, or inorganic–organic compounds are bonded through strong covalent or ionocovalent chemical bonds. In some cases, classes I and II are, in fact, mixed together.

5.1 Nanocomposites Containing Silica Phases

Mark (1996) employed several ways to incorporate silica-type fillers *in situ* into polydimethylsiloxane (PDMS) networks. The linear PDMS chains used were silanol-terminated or vinyl-terminated. The techniques employed were: (i) incorporating silica that has been precipitated *in situ* from a silicate in an aqueous dipersion with PDMS; (ii) precipitating silica directly into PDMS during its curing; (iii) precipitating silica into a swollen PDMS network after it has been cured; (iv) incorporating silica prepared from TEOS and containing some PDMS; and (v) incorporating silica prepared from partially hydrolyzed TEOS and also containing some PDMS. The morphology of the silica particles generated *in situ* strongly depends on the catalysts used. Good definition of nano-silica particles with a narrow size distribution generally occurs when the catalyst is a base, such as ethylamine or diethylamine. On the other hand, use of an acidic catalyst gives poorly defined "fuzzy" particles. This lack of definition is consistent with results in the sol–gel ceramics area, where it was concluded that acidic catalysts give structures that are less branched and less compact than those obtained using basic catalysts. These morphologies were observed by transmission electron microscopy (TEM). Moreover, at some compositions and under some hydrolysis conditions that included an acid, bicontinuous silica and PDMS phases or the silica and PDMS phases interpenetrating one another can be obtained. This was evidenced by small-angle x-ray scattering and small-angle neutron scattering. Another interesting phenomenon was also observed: under basic conditions silica particles remaining in contact with their aqueous catalyst solution can "age" or "digest." As a result, the particles became better defined and were more uniform in size. There seem to be interesting parallels with Ostwald ripening in the area of colloid science. In these *in situ*-precipitated silica systems, significant reinforcements have been obtained for PDMS networks under various deformations (see *Elastomers, Ceramic-modified** and *Polymer–Silica Nanocomposites**).

5.2 Nanocomposites Containing Other Ceramic Phases

In addition to silica, Mark (1996) also showed that some nonsilica fillers can also provide good reinforcement for PDMS networks. These include titania (TiO_2), alumina (Al_2O_3), and zirconia (ZrO_2). Using these nonsilica fillers can eliminate a problem associated with acidic hydroxysilyl groups on the silica surface that can cause PDMS chain cleavage at high temperatures. In the case of titania-reinforced PDMS systems it was observed, interestingly, that the stress–strain isotherms have much better reversibility.

Ceramics are often bonded to oligomers or polymers as modifiers. The majority of Wilkes' work (Wen and Wilkes 1996) has emphasized the oligomeric or polymeric species as the dominant component within the final network material. Hence, these materials tend to be flexible and have been called "ceramers" to indicate some combination of ceramic-like behavior with polymeric-like behavior.

Wen and Wilkes (1996) reported structure–property relationships for organic–inorganic hybrid systems prepared by the sol–gel process using either titanium isopropoxide or zirconium isopropoxide with silane end-capped poly(tetramethylene oxide) (PTMO) oligomers. These novel network materials can be monolithic and are optically transparent, as were ceramers prepared with PTMO and TEOS as the metal alkoxide. The dynamic mechanical behavior of the titanium- and zirconium-containing ceramers exhibits trends similar to those reported for the TEOS/PTMO ceramers. However, the stress–strain response shows yielding behavior and higher moduli and ultimate stress at break. Similarly, metal acetyl acetonates of aluminum and zinc were also incorporated into PTMO-modified TEOS based materials.

Clays are interesting as ceramic phases since polymers can be inserted into the interlayer spacing of some of these minerals, either expanding them or even exfoliating them (Okada *et al.* 1990). Relatively little clay is required to give large improvements in mechanical properties. The sheet-like structure of the dispersed phases is also useful for decreasing the permeability of the host polymer, which is important in applications such as fuel tanks (see *Clay-based Polymer Nanocomposites** and *Polymer–Nonsilica Ceramic Nanocomposites**).

5.3 Nanocomposites Containing Mixed Ceramic Phases

In many cases a single ceramic component cannot produce the desired properties. Therefore, two or more metal alkoxides are often used together, with balanced reaction rates, to tailor the desired properties. The hard contact lens material developed by Schmidt (Schmidt and Philipp 1984) is a good example. The structure of the material is illustrated in Fig. 1. In this system, an epoxide-substituted alkoxysilane and titatium alkoxides were used as precursors for the hydrolysis and condensation reactions. Titanium alkoxide was used to catalyze the silanol condensation. To achieve sufficient tensile strength, polymethacrylates were incorporated by using a methacryloxy-substituted alkoxysilane as a link between the siliceous network and the polymer chain. The incorporation of titania led to dense monolithic products that could be cured with only some minor shrinkage. Good wettabilities were due to glycol groups formed by the epoxide radicals. The

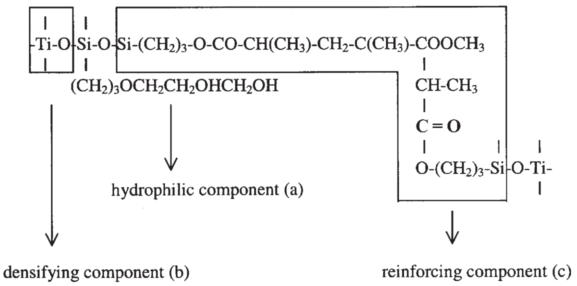

-Ti-O-Si-O-Si-(CH2)3-O-CO-CH(CH3)-CH2-C(CH3)-COOCH3
(CH2)3OCH2CH2OHCH2OH
CH-CH3
C=O
O-(CH2)3-Si-O-Ti-
hydrophilic component (a)
densifying component (b)
reinforcing component (c)

Figure 1
A hybrid inorganic–organic nanocomposite for use as a new contact lens material prepared by the sol–gel process based on the principle of molecular engineering (Philipp and Schmidt 1984). Molar ratio a:b:c = 65:5:30.

excellent oxygen permeability resulted from the siloxane structure. It can be seen that this hybrid inorganic–organic composite is a product based on a principle of molecular engineering.

This brief review shows that creation of an appropriate ceramic phase and achievements of useful functions depend on an appropriate choice of metal alkoxide. Similarly, the morphological structure of the ceramic phase is strongly related to the hydrolysis and condensation reactions. Through the combination of polymers and ceramics, new materials can be created and their properties can be tailored to a wide variety of applications.

Bibliography

Bradley D C, Mehrotra R C, Gaur D P 1978 *Metal Alkoxides*. Academic Press, New York

Brinker C J, Scherrer G W 1990 *Sol-Gel Science. The Physics and Chemistry of Sol-Gel Processing*. Academic Press, San Diego, CA

Chujo Y, Saegusa T 1992 Organic polymer hybrids with silica gel formed by means of the sol–gel method. *Adv. Polym. Sci.* **100**, 11–29

Guglielmi M, Carturan G 1988 Precursors for sol–gel preparations. *J. Non-Cryst. Solids* **100**, 16–30

Komarneni S, Parker J C, Wollenberger H J (eds.) 1997 *Nanophase and Nanocomposite Materials II*. Materials Research Society, Pittsburgh, PA

Mark J E 1996 The sol–gel route to inorganic–organic composites. *Heterog. Chem. Rev.* **3**, 307–20

Mark J E, Lee C Y-C, Bianconi P A (eds.) 1995 *Hybrid Organic–Inorganic Composites*. American Chemical Society, Washington, DC

Novak B 1993 Hybrid nanocomposite materials between inorganic glasses and organic polymers. *Adv. Mater.* **5**, 422–33

Okada A, Kawasumi M, Usuki A, Kojima Y, Kurauchi T, Kamigaito O 1990 Nylon 6–clay hybrid. In: Schaefer D W, Mark J E (eds.) *Polymer Based Molecular Composites*. Materials Research Society, Pittsburgh, PA, pp. 45–50

Philipp G, Schmidt H 1984 New materials for contact lenses prepared from Si- and Ti-alkoxides by the sol–gel process. *J. Non-Cryst. Solids* **63**, 283–92

Sanchez C, Ribert F 1994 Design of hybrid organic–inorganic materials synthesized via sol–gel chemistry. *New J. Chem.* **18**, 1007–47

Schaefer D W, Mark J E (eds.) 1990 *Polymer Based Molecular Composites*. Materials Research Society, Pittsburgh, PA

Schmidt H 1994 Inorganic–organic composites by sol–gel techniques. *J. Sol-Gel Sci. Tech.* **1**, 217–31

Schubert U, Husing N, Lorenz A 1995 Hybrid inorganic–organic materials by sol–gel processing of organofunctional metal alkoxides. *Chem. Mater.* **7**, 2010–27

Sharp K G 1998 Inorganic/organic hybrid materials. *Adv. Mater.* **10**, 1243–8

Wen J, Wilkes G L 1996 Organic/inorganic hybrid network materials by the sol–gel approach. *Chem. Mater.* **8**, 1667–81

Yoldas B E 1975 Alumina gels that form porous transparent Al_2O_3. *J. Mater. Sci.* **10**, 1856–60

P. Xu
W. L. Gore & Associates, Inc., Elkton, Maryland, USA

Polymer–Ceramic Nanocomposites: Control of Shrinkage

One of the most complex aspects of sol–gel processing of polymer–ceramic nanocomposites is the drying process. Because of the cosolvents, liberated alcohol, and excess water used, volume shrinkages of 80% or more are common. Fibers and thin films can be dried fairly quickly and easily, but thick films and monoliths must be dried under carefully controlled conditions to avoid the formation of cracks, deformations, and residual stresses in the material. For a detailed account of the drying process see *Xerogels**. The homogeneous incorporation of organic polymers into the ceramic phase adds even more complexity to the problem. Not only must the drying rate be controlled to avoid mechanical failure of the sample, controlled drying is also necessary to prevent phase separation of the organic polymer from the inorganic matrix. Furthermore, the large volume shrinkage associated with drying precludes the use of standard forming processes such as extrusion, molding, or casting.

For these reasons, several important and useful methods have been developed to reduce or eliminate the shrinkage associated with sol–gel processing of polymer–ceramic nanocomposites. A typical sol–gel derived polymer–ceramic nanocomposite is prepared by dissolving the desired polymer in an organic solution with tetraethylorthosilicate (TEOS), water, and an acidic or basic sol–gel catalyst. After gelation of the silica network the solvent, water, and ethanol liberated during the hydrolysis and condensation of TEOS are typically removed by slow evaporation. This process can be altered in several different ways to control shrinkage. A two-step sol–gel process can be used where the solvent is evaporated or distilled prior to gelation, so that after gelation less solvent will need to be removed, resulting in less shrinkage. Alternatively, the solvent can be displaced after gelation, but prior to drying, with a polymerizable solvent to produce solid monoliths or with a supercritical solvent to produce porous materials called aerogels (for more details of this process see *Organic Aerogels** and *Inorganic and Inorganic–Organic Aerogels**). In a more direct process, nonshrinking sol–gel composites can be prepared from ceramic precursors with polymerizable alkoxides and a polymerizable solvent so that all of the components in the solution are converted to either ceramic or polymer in a single step. Finally, preformed metal oxide oligomers and colloids can be suspended in a polymerizable solvent followed by subsequent polymerization of the solvent, which also converts all of the components to either ceramic or polymer.

1. Two-step Sol–Gel Process

In the two-step or two-stage sol–gel process, the hydrolysis and condensation step and the gelation step are accomplished in two separate and distinct operations rather than occurring simultaneously, as is the case in standard sol–gel processing. The first step in the process is to completely hydrolyze and partially condense the TEOS. This is accomplished by heating an acidic solution of TEOS in a water-miscible solvent such as ethanol or tetrahydrofuran (THF) with the polymer or polymer precursor in the solution. A portion of the solvent is removed by distillation, resulting in a viscous solution. In the second step, this solution is cast into a film or mold and allowed to gel either directly from the acidic solution, or the acidic solution can be rendered basic prior to casting to promote faster gelation. Since some solvent remains in the system, drying is still necessary and shrinkage is observed. However, since a portion of the solvent has been removed, the amount of shrinkage is reduced, resulting in less stress in the material during the final drying step.

Wei *et al.* (1996) describe a convenient technique for the preparation of transparent vinyl polymer–silica hybrid sol–gel composites with less than 20% volume shrinkage (Fig. 1). The first step in the reaction is the hydrolysis of TEOS in the presence of 2-hydroxyethyl methacrylate (HEMA) in THF solvent. A vinyl polymerization initiator such as benzoyl peroxide is added to the system and the solvent, water, and ethanol removed by vacuum stripping, resulting in a clear, viscous liquid. At this stage other reactive monomers, polymers, or silane coupling agents can be added to the system. The homogeneous fluid is then cast into molds followed by initiation of the vinyl polymerization and further formation of the inorganic network at elevated temperatures. Composites containing up to 50 wt.% silica were prepared with less than 20% volume shrinkage. Higher silica content composites display higher percentage volume shrinkage, but in all cases the volume shrinkage is less than that observed using conventional sol–gel techniques. This same approach was also applied to the preparation of materials for multimode channel waveguides (Xu *et al.* 1996).

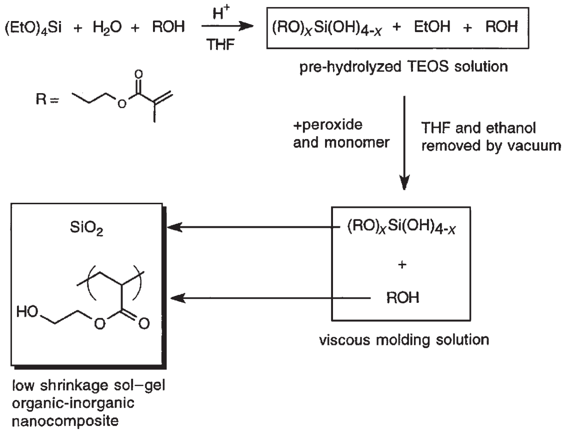

$(EtO)_4Si + H_2O + ROH \xrightarrow[THF]{H^+} (RO)_xSi(OH)_{4-x} + EtOH + ROH$
pre-hydrolyzed TEOS solution
R =
+peroxide and monomer
THF and ethanol removed by vacuum
$(RO)_xSi(OH)_{4-x}$ + ROH
viscous molding solution
SiO_2
HO
low shrinkage sol–gel organic-inorganic nanocomposite

Figure 1
Reaction sequence for the preparation of low shrinkage polymer–ceramic nanocomposites by a two-step sol–gel process.

Using a two-step process, epoxy/poly(methyl methacrylate) (PMMA)/silica hybrids were prepared and evaluated for contact lens use (Phillip and Schmidt 1984). Composites were prepared by combining TEOS, methacryloxypropyl trimethoxysilane, an epoxysilane coupling agent, and methyl methacrylate (MMA) or HEMA. When the silica phase was sufficiently formed, the viscous solution was poured into a mold and cured at 150 °C to initiate the MMA polymerization and complete the silica network formation. The use of silane coupling agents in the process prevents phase separation during the drying process and also improves the mechanical properties of the composites.

2. Solvent Displacement

One method for eliminating the shrinkages associated with solvent evaporation is to replace the solvent with a reactive liquid which can be directly converted to a polymer (Fig. 2). The advantage of this approach is that the structure and morphology of the silica phase is controlled by the sol–gel processing conditions and no special adjustments of the process are required to accommodate the polymer during the formation of the silica network. An important limitation to this approach is that access to the silica gel monolith

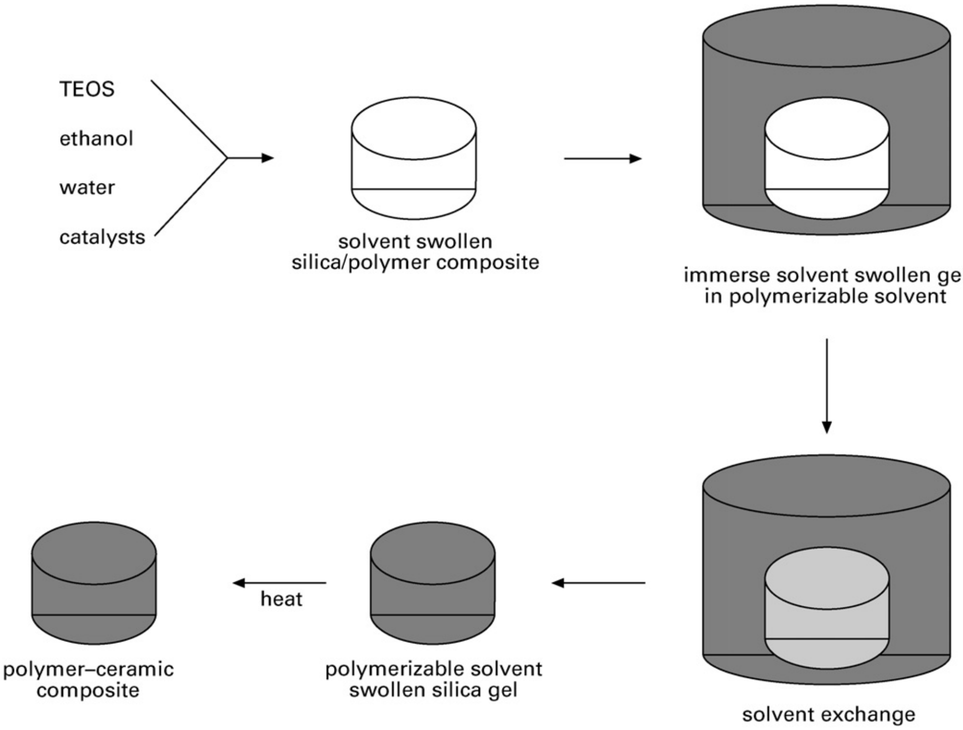

TEOS
ethanol
water
catalysts
solvent swollen silica/polymer composite
immerse solvent swollen gel in polymerizable solvent
solvent exchange
polymerizable solvent swollen silica gel
heat
polymer–ceramic composite

Figure 2
The solvent exchange sequence for producing polymer–ceramic composites.

surfaces is required to effect the solvent exchange. This requires either a mold cavity with at least one open face, or handling of the very fragile solvent-swollen gel. Furthermore, achieving high ceramic ratios requires partial drying of the gel. The process of partial drying is complicated by the fact that some level of densification must be achieved in order to maintain the same network structure once the reactive solvent is introduced. If no further inorganic network formation has occurred in the partially dried gel, the material can swell to its original volume when reintroduced to a solvent. A final, but minor, limitation is that the polymerization of the reactive solvent is limited to addition polymerization since condensation polymerization would result in volatile by-products which cause unwanted shrinkage and porosity.

Pope *et al*. (1989) prepared sol–gel derived silica from TEOS in ethanol and then placed the solvent-swollen silica gel monolith into MMA. An equilibrium solution is established between the ethanol in the gel and the MMA resulting in diffusion of the ethanol out of the gel and infusion of the MMA into the silica gel matrix. By repeatedly decanting the MMA/ethanol solution from the silica gel monolith and adding fresh MMA, eventually all of the ethanol is removed. A free radical initiator such as a peroxide or azo compound is added to the system, followed by outgassing and heating to initiate the polymerization. After the polymerization is complete, excess PMMA is removed by polishing to obtain the finished silica–PMMA composite. The ratio of silica to PMMA is increased by partially drying the solvent-swollen silica gel prior to solvent exchange with MMA. In an extension of this approach, the silica glass was treated with methacryloxypropyl trimethoxysilane to provide a chemical bond between the PMMA and the glass.

In cases where the solvent used in the sol–gel process and the reactive solvent are not miscible, a secondary solvent exchange with a more general solvent, such as THF, is performed to replace the sol–gel solvent with a solvent that is miscible with the desired reactive solvent. If a porous composite is desired, the sol–gel solvent is exchanged with liquid carbon dioxide under supercritical conditions to produce composite aerogels (see *Organic Aerogels** and *Inorganic and Inorganic–Organic Aerogels**).

3. *Reactive Sol–Gel Precursors*

The two major components that are removed during the drying process of sol–gel derived composites are the solvent and the ethanol liberated during the hydrolysis and condensation of TEOS. By using a

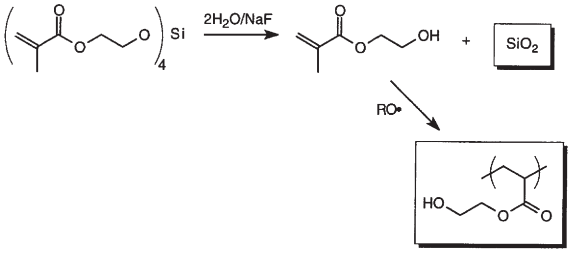


Figure 3
Preparation of nonshrinking sol–gel polymer–ceramic nanocomposites from a silicone tetraalkoxide bearing polymerizable alkoxide groups.

tetraalkoxysilane derivative with polymerizable alkoxides, a polymerizable cosolvent and a stoichiometric quantity of water, all components in the reaction solution are converted to silica or polymer. Because all of the components are converted to solids, drying is unnecessary and the only shrinkage observed is the minor shrinkage associated with the bulk organic polymerization (Fig. 3). Because all of the components are converted to either ceramic or organic polymer in one step, processes such as reaction injection molding are possible, allowing for direct fabrication of monolithic composites and complex shapes.

This ‘‘nonshrinking’’ sol–gel composite process has been applied to silica-based systems using ring-opening metathesis polymerization (ROMP) (Ellsworth and Novak 1991) and free radical polymerization (Novak and Davies 1991). In order to achieve complete formation of both the inorganic network and the organic polymer network, the reaction rates of both the sol–gel polymerization and the organic polymerization must be carefully matched. If the inorganic network is allowed to form at a much faster rate than the organic polymerization, premature termination of the organic polymerization results, causing shrinkage in the composite due to the evaporation of unreacted monomer. This effect is most significant in transition metal catalyzed polymerizations such as ROMP. More importantly, the organic polymerization must not proceed faster than the inorganic network formation. In this case, large-scale phase separation of the polymer occurs and the inorganic network can no longer form in the solid organic polymer phase. Under ideal reaction conditions transparent polymer–ceramic nanocomposites are obtained, with no drying step needed.

4. *Preformed Ceramic Oligomers and Colloids in Polymerizable Solvents*

An important limitation to the polymerizable silicone tetraalkoxide approach described above is that the maximum inorganic content is limited to 10–18 wt.% by the internal stoichiometry of the silica precursor (4 equivalents of polymer for every 1 equivalent of SiO_2). The glass content of nonshrinking sol–gel composites is increased by employing highly branched precondensed silica precursors such as poly(silicic acid) or colloidal silica. Poly(silicic acid) is prepared by the hydrolysis of sodium metasilicate in acidic aqueous solution resulting in a highly branched silica oligomer. The Stober process, which is used to prepare colloidal silica, involves the hydrolysis and condensation of TEOS in basic ammonium hydroxide/ethanol solution. Either form of silica is reacted with a polymerizable alcohol in an esterification reaction to produce an organically soluble, high silica content precursor for nonshrinking sol–gel composite preparation under the same reaction conditions employed in the polymerizable silicone tetraalkoxide approach. In both cases, all components in the reaction contribute to the inorganic network or the organic polymer network, but with higher inorganic content composites obtainable by using a preformed silica oligomer. Preformed ceramics such as colloidal silica and fumed silica have been used for many decades in the production of polymer–ceramic composites where only the polymer phase is continuous. The distinction in the systems described here is the simultaneous formation of an inorganic network within the polymer network resulting in levels of dispersion and stability not obtainable from conventional composites.

Poly(silicic acid) oligomers of varying molecular weight and degrees of branching are prepared by reacting sodium metasilicate in acidic aqueous solution. The molecular weight of the oligomers is controlled by increasing the reaction time to achieve higher molecular weights. By increasing the pH of the reaction solution, higher degrees of branching are achieved. Poly(silicic acid) esters are prepared from poly(silicic acid) and various polymerizable alcohols by a modified esterification process. The poly(silicic ester)s of acrylate derivatives in particular must be prepared under carefully controlled conditions to prevent spontaneous polymerization of the acrylate monomer. The degree of esterification must be high enough (greater than 50% of the total available reactive sites) for proper solubility of the poly(silicic ester) in a polymerizable cosolvent. Because the poly(silicic acid ester) still contains reactive silicone hydroxyl sites, the ester must be stored in dilute solution. The poly(silicic acid ester) is then used for nanocomposite preparation. For example, a poly(HEMA)–silica nanocomposite containing 50 wt.% silica was prepared by dissolving the HEMA ester of poly(silicic acid) in HEMA with benzoyl peroxide as a free radical initiator. A transparent nonshrinking composite was obtained after heating the sample under nitrogen to complete the formation of the inorganic network and to initiate the free radical polymerization. This process was applied to several

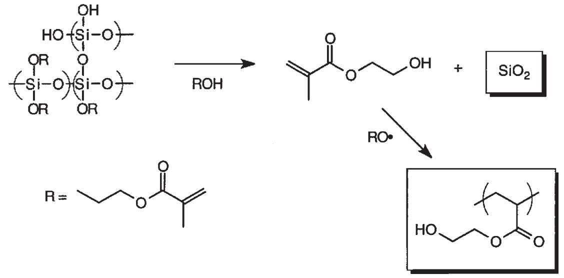


Figure 4
Sequence for preparing high glass content nonshrinking sol–gel polymer–ceramic nanocomposites from poly(silicic acid ester)s.

different acrylate esters as well as an oxanorbornene–poly(silicic acid) ester polymerized by a transition metal ROMP catalyst (Fig. 4).

Bibliography

Ellsworth M W, Novak B M 1991 Mutually interpenetrating inorganic–organic networks. New routes into nonshrinking sol–gel composite materials. *J. Am. Chem. Soc.* **113**, 2756–8

Ellsworth M W, Novak B M 1993 "Inverse" organic–inorganic composite materials. 3. High glass content "nonshrinking" sol–gel composites via poly(silicic acid esters). *Chem. Mater.* **5**, 839–44

Novak B M 1993 Hybrid nanocomposite materials—between inorganic glasses and organic polymers. *Adv. Mater.* **5**, 422–33

Novak B M, Davies C 1991 "Inverse" organic–inorganic composite materials. 2. Free-radical routes into nonshrinking sol–gel composites. *Macromolecules* **24**, 5481–3

Phillipp G, Schmidt H 1984 New materials for contact lenses prepared from Si- and Ti- alkoxides by the sol–gel process. *J. Non-Cryst. Solids* **63**, 283–92

Pope E J A, Asami M, Mackenzie J D 1989 Transparent silica gel–PMMA composites. *J. Mater. Res.* **4**, 1018–26

Pope E J A, Mackenzie J D 1986 Porous and dense composites from sol–gel. *Mater. Sci. Res.* **20**, 187–94

Wei Y, Jin D, Yang C, Wei G 1996 A fast convenient method to prepare hybrid sol–gel materials with low volume-shrinkages. *J. Sol-Gel Sci. Technol.* **7**, 191–201

Xu C, Eldada L, Wu C, Norwood R A, Shacklette L W, Yardley J T, Wei Y 1996 Photoimageable, low-shrinkage organic–inorganic hybrid materials for practical multimode channel waveguides. *Chem. Mater.* **8**, 2701–3

M. W. Ellsworth
Tyco Electronics, Menlo Park, California, USA

Polymer–Ceramic Nanocomposites: Interfacial Bonding Agents

Organic–inorganic composites represent a special class of materials with properties intermediate between organic polymers and inorganic ceramers (see *Polymer–Nonsilica Ceramic Nanocomposites**; *Polymer–Silica Nanocomposites**). The primary goal of preparing such composites is to achieve combinations of properties from both components. One successful approach for preparing organic–inorganic composites has been the *in situ* polymerization of metal alkoxides within organic polymer matrices via sol–gel process. This process is based on hydrolysis of metal alkoxides to produce metal hydroxides followed by polycondensation of the hydroxyl groups to form a three-dimensional network.

Unfortunately, incompatibility of two phases most often tends to result in undesirable phase separation. Thus, the interactions between the organic phase and the inorganic phase is paramount in constraining phase separation. The possible interactions between the organic phase and the inorganic phase can be achieved through hydrogen bonding, employing organic polymers with hydroxy or alkoxy terminating chains which can participate in the hydrolysis reaction used to generate ceramic, or bonding the dissimilar phases with the use of a bonding agent (O'Reilly and Coltrain 1996).

1. Interfacial Bonding

Bonding of organic polymers to inorganic surfaces with the primary goal of preparing composites with high strength-to-weight ratio has long been known (Plueddemann 1991). In this approach, two dissimilar phases are held together by use of a coupling agent, also known as a bonding agent. This bonding agent, in general, can undergo hydrolysis with the sol–gel precursor forming the inorganic network, at the same time it can react with the organic polymer. Improved interfacial bonding between phases retards phase separation, giving materials with enhanced properties. Thus, the overall quality of the composites can be improved by the use of bonding agents, especially under high-humidity and high-temperature conditions. Silanes are most widely used as coupling agents (Plueddemann 1991). However, other coupling agents such as organotitanates, phosphorous-containing compounds, and chromium coordination complexes are also used.

2. Silane Coupling Agents

Silane coupling agents play an important role in the preparation of composites from organic polymers and inorganic fillers such as glass, minerals, and metals. The general structure of silane coupling agents are $(RO)_3$-Si-R‘-X, where X is an organofunctional group, R‘ is a small alkylene linkage, and RO is a hydrolyzable group such as an alkoxy group (Pape 1996). The organofunctional group and the alkoxy group can form strong bonds with the polymer matrix and with the filler particles, respectively. The number of hydroxy groups resulting from the hydrolysis of alkoxy groups can be controlled by adjusting the rates of hydrolysis and condensation of the silane coupling agent to achieve optimal bonding effects (Premachandra *et al.* 1998). Figure 1 illustrates the mechanism of a silane coupling agent to improve the interfacial bonding between two phases.

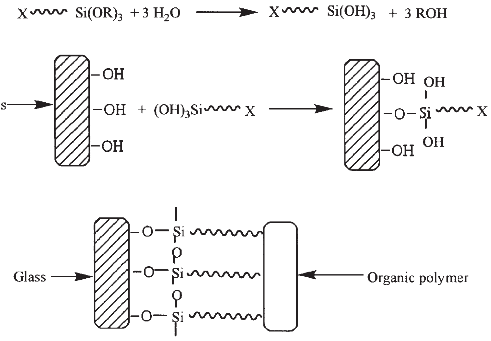


Figure 1
Improved interfacial bonding by a silane coupling agent.

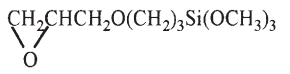
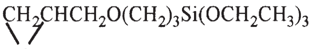
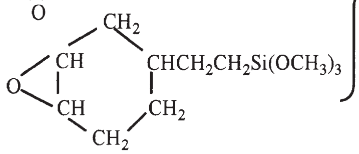
Organic groups such as vinyl, allyl, amino, methacryloxy, and epoxy groups, in general, react favorably with polymers (Stepek and Daoust 1983). Some commercially available silane coupling agents which can be utilized for different polymer systems are given in Table 1 (Stepek and Daoust 1983, Pape 1996).

3. Bonding in High-temperature Polymer–ceramic Composites

High-performance high-temperature polymers, in general, are difficult to treat in the usual sol–gel technique (see *Ceramic-modified High-temperature Polymers*). Because of their intractability, the processing is difficult, and their unreactivity often causes poor bonding between the organic polymer and the inorganic ceramic phase, resulting in undesirable phase separation. Therefore, improvements in bonding between the phases is critical in high-temperature polymers to achieve the desired properties.

Coupling agents used for high-temperature polymers should be of high thermal stability comparable with the polymer. Aromatic silanes, in general, have much better thermal stability and oxidative resistance

Table 1
Some commercially available silane coupling agents for different polymer systems.

Coupling agent	Structure	Polymers
γ-Aminopropyltrimethoxysilane	$NH_2CH_2CH_2CH_2Si(OCH_3)_3$	PE, PP, PMMA, PC, PA, PUR, Pol sulfones, Polysulfides, Epoxies, Melamine resins, Phenoplasts
N-β-(Aminoethyl)-γ-aminopropyltrimethoxysilane	$NH_2CH_2CH_2NH(CH_2)_3Si(OCH_3)_3$	
γ-Glycidoxypropyltrimethoxysilane	 $CH_2CHCH_2O(CH_2)_3Si(OCH_3)_3$	PE, PP, PA, PC, PS, ABS, PVC, PUR, Polysulfides, Polyesters, Epoxies, Phenoplasts, Melamine resins
γ-Glycidoxypropyltriethoxysilane	 $CH_2CHCH_2O(CH_2)_3Si(OCH_2CH_3)_3$	
β-(3,4-Epoxycyclohexyl)ethyltrimethoxysilane		
γ-Mercaptopropyltrimethoxysilane	$HSCH_2CH_2CH_2Si(OCH_3)_3$	PS, PVC, PPO, Polysulfones, Polysufides EPDM, SBR, NR, PUR, Polymercaptans, Epoxies
β-Mercaptoethyltriethoxysilane	$HSCH_2CH_2Si(OC_2H_5)_3$	
Vinyltriethoxysilane	$CH_2{=}CHSi(OC_2H_5)_3$	PE, PP, PVC, SBR, EPM, EPDM, Epoxies
Vinyl-tris(β-methoxyethoxy)silane	$CH_2{=}CHSi(OCH_2CH_2OCH_3)_3$	

than aliphatic silanes. However, the aromatic silanes lack the high degree of reactivity which is required to function as coupling agents. High-temperature polymer composites have been prepared using some high-temperature polymers, including aromatic polyamides (Wang *et al.* 1993, Ahmad *et al.* 1997), polyimides (Morikawa *et al.* 1992, Wang *et al.* 1994), and some poly(benzoxazoles) (Premachandra *et al.* 1996), and poly(benzobisthiazoles) (Premachandra *et al.* 1996, Premachandra *et al.* 1999), with proper bonding between phases achieved by use of bonding agents.

3.1 Polyimide–ceramic Composites

Polyimide-silica hybrid materials prepared without using a bonding agent were transparent only up to a 8 wt.% of silica content (Morikawa *et al.* 1992). Scanning electron microscopy results demonstrated silica particles with diameters of approximately 5 μm in such films. However, increasing the chemical bonding between phases by using a bonding agent, specifically aminophenyltrimethoxysilane (APTMOS), increased dramatically the transparency of the materials was and the silica particle sizes were reduced (Wang *et al.* 1994). Thermal analysis showed no apparent decrease in thermal decomposition temperature by using APTMOS. Polyimide–silica hybrid materials prepared without a bonding agent showed an increase in the modulus but a decrease in the tensile strength. However, the improved bonding between the two phases increased both the modulus and tensile strength. The hardness measurements of hybrid materials with bonding between phases also showed an increase relative to the pure polymer whereas the hardness in hybrid materials prepared without the bonding agent decreased, as was the case for the tensile strength.

3.2 Polyamide–ceramic Composites

Poly(phenylene terephthalamide) have been used to prepare composites, in some instances, with improved interfacial bonding achieved by bonding agents. The composites prepared with the bonding agent APTMOS showed enhancements in the tensile strength and modulus with increases in the amount of the inorganic phase, silica (Wang *et al.* 1993, Ahmad *et al.* 1997). Composites prepared without the bonding agent, however, showed decreases in the tensile strength on incorporation of filler (Ahmad *et al.* 1997). Using a nonlinear polyamide with a large number of chain ends able to react with the inorganic network could increase the binding between two phases. The improvements in the tensile strength in this case were significant compared to the pure polyamide, which, however, showed a decrease in the tensile strength compared to the corresponding linear polyamide. The transparencies in both systems, linear and branched polyamides, in which the two phases were chemically bonded were improved compared to those composites prepared with no bonding agent.

3.3 Poly(benzoxazole)–ceramic Composites

Poly(benzoxazoles) (PBOs), in general, are intractable and insoluble in most organic solvents, causing difficulties when fabricating them into useful objects. By introducing specific functional groups and preparing functionalized copolymers, these polymers can be made somewhat more tractable. Hydroxy-functionalized copolymers of PBOs (HPBOs) which are soluble in tetrahydrofuran (THF) have been used to prepare composites using silica as the inorganic phase (Premachandra *et al.* 1996). HPBO–silica composites prepared without bonding could be redissolved in THF but the composites prepared with a bonding agent, specifically isocyanatopropyltriethoxysilane, were insoluble in THF. In addition, the transparencies of hybrid films were improved compared to those without the bonding agent. Substantial improvements in the tensile modulus and tensile strength for the composites could be achieved when a bonding agent is used.

3.4 Poly(benzothiazole)–ceramic Composites

Highly transparent hybrid materials prepared from sulfo-functionalized poly(benzothiazoles) (SPBTs) and silica have been reported (Premachandra *et al.* 1996). The insolubility of the resulting films in methanol suggests sufficient bonding between the polymer and silica phase. Substantial increases in mechanical and thermal properties were exhibited by the SPBT–silica composites compared to the pure polymer.

4. Summary

The improved interfacial bonding between phases can be achieved by using suitable bonding agents in the preparation of organic–inorganic composites. Thus, the macroscopic phase separation which is often a problem in the preparation of such composites can be suppressed. Improved interfacial bonding exerts a profound effect on the morphology of the resulting hybrid materials: it tends to suppress the agglomeration of ceramic particles, resulting in reduced particle sizes, thus increasing the surface area which can interact with the polymer. This allows the ceramic particles to interact with the polymer backbone at a molecular level, resulting in substantial reinforcement of the hybrid material.

See also: Ceramic-modified High-temperature Polymers*; Polymer–Nonsilica Ceramic Nanocomposites*; Polymer–Silica Nanocomposites*

Bibliography

Ahmad Z, Sarwar M I, Mark J E 1997 Chemically bonded silica–polymer composites from linear and branched polyamides in a sol–gel process. *J. Mater. Chem.* **7** (2), 259–63

Morikawa A, Iyoku Y, Kakimoto M, Imai Y 1992 Preparation of a new class of polyimide-silica hybrid films by sol–gel process. *Polym. J.* **24**, 107–13

O'Reilly J M, Coltrain B K 1996 Organic–inorganic composite materials. In: Salamone J C (ed.) *Polymeric Materials Encyclopedia: Synthesis, Properties and Applications*. CRC Press, Boca Raton, FL

Pape P G 1996 Silane coupling agents. In: Salamone J C (ed.) *Polymeric Materials Encyclopedia: Synthesis, Properties and Applications*. CRC Press, Boca Raton, FL

Plueddemann E P 1991 *Silane Coupling Agents*. Plenum, New York

Premachandra J, Kumudinie C, Zhao W, Mark J E, Dang T D, Chen J P, Arnold F E 1996 Polymer–silica hybrid materials prepared from some functionalized polybenzoxazoles and polybenzobisthiazoles. *J. Sol–gel Sci. Technol.* **7**, 163–75

Premachandra J K, Kumudinie C, Mark J E, Dang T D, Arnold F E 1999 Preparation and properties of some hybrid aerogels from a sulfopolybenzobisthiazole-silica composite. *J. Macromol. Sci.-Pure Appl. Chem.* **A36** (1), 73–83

Premachandra J K, van Ooij W J, Mark J E 1998 Reaction kinetics of γ-ureidopropyl-trimethoxysilane in the water–methanol system by FTIR spectroscopy. *J. Adhesion Sci. Technol.* **12**, 1361–76

Stepek J, Daoust H 1983 *Additives for Plastics*. Springer, New York

Wang S, Ahmad Z, Mark J E 1993 A polyamide–silica composite prepared by the sol–gel process. *Polym. Bull.* **31**, 323–30

Wang S, Ahmad Z, Mark J E 1994 Polyimide–silica hybrid materials modified by incorporation of an organically substituted alkoxysilane. *Chem. Mater.* **6**, 943–6

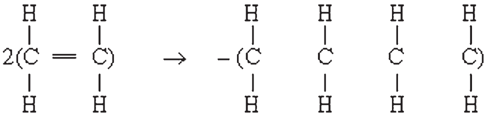
C. Kumudinie
University of Cincinnati, Ohio, USA

Polymer–Ceramic Nanocomposites: Polymer Overview

Macroscopic inorganic–organic hybrids, notably fiberglass-reinforced epoxy, are well accepted composite materials. While they are essential to many industries, such as electronics, aerospace, and automotive, macroscopic composites have lost some of their glamour. Now the excitement in advanced materials is nanocomposites. Nanocomposites refer to the family of materials where the inorganic and organic components are mixed on the nanometer scale. This first section introduces the polymers that have been mixed in various ways with inorganic materials, mainly zeolites, clays, ceramics, and glasses. The details of synthesis and the discussions of structure–property relationships are left to the sections that follow.

This introduction includes both (i) simple textbook definitions of terms used in polymer technology and (ii) a "roadmap" to the literature on nanocomposites. Principal references and source materials in this field include special issues of the *Journal of Sol–Gel Science and Technology* (1995, 1996), Drzal *et al.* (1995), Coltrain *et al.* (1996), Laine *et al.* (1998), Klein *et al.* (1999) and Andrews and Najafi (1997). Also, see *Polymer–Silica Nanocomposites**.

By way of definition, a polymer is a long-chain molecule, a macromolecule, usually with a carbon backbone. The building block is the monomer. The polymer increases in molecular weight either by (i) addition polymerization or (ii) condensation polymerization. Addition polymerization involves the reaction of breaking a double bond to form two single bonds. A typical addition polymerization for ethylene is:



$$2\left(\begin{array}{c}\mathrm{H}\\|\\\mathrm{C}\\|\\\mathrm{H}\end{array} = \begin{array}{c}\mathrm{H}\\|\\\mathrm{C}\\|\\\mathrm{H}\end{array}\right) \rightarrow -\left(\begin{array}{c}\mathrm{H}\\|\\\mathrm{C}\\|\\\mathrm{H}\end{array}\quad\begin{array}{c}\mathrm{H}\\|\\\mathrm{C}\\|\\\mathrm{H}\end{array}\quad\begin{array}{c}\mathrm{H}\\|\\\mathrm{C}\\|\\\mathrm{H}\end{array}\quad\begin{array}{c}\mathrm{H}\\|\\\mathrm{C}\\|\\\mathrm{H}\end{array}\right)- \qquad (1)$$

with the application of heat or light. Typical addition polymers are polyethylene, poly(vinyl chloride) (PVC), poly(vinyl acetate) (PVAc), poly(methyl methacrylate) (PMMA), and polystyrene.

Condensation polymerization involves the reaction of two small molecules and produces a nonpolymerizable by-product. A polymerization for nylon is:

$$\underset{\text{Aminoacid}}{2\mathrm{NHRR'COOH}} \rightarrow \underset{\text{“B”}}{\mathrm{NHRR'CO{-}NRR'COOH}} + \mathrm{H_2O} \qquad (2a)$$

$$\text{“B”} + \mathrm{NHRR'COOH} \rightarrow$$

$$\mathrm{NHRR'CO{-}NRR'CO{-}NRR'COOH} + \mathrm{H_2O} \qquad (2b)$$

where R is a univalent radical and R$'$ is a divalent radical ($(\mathrm{CH_2})_n$ where $n \geqslant 5$). The first step is carried out at 175 °C with a catalyst and the second step is carried out at 250 °C under vacuum. Typical condensation polymers include nylon (hexamethylamine and adipic acid), polyamides, polyimides, and polyesters.

The functionality of the polymer refers to the number of reaction sites available. A bifunctional monomer produces linear polymers. A trifunctional or tetrafunctional monomer produces network polymers. When polymerization occurs between two different monomers, the result is a copolymer.

A reasonable question to ask at this point is why combine organic and inorganic materials? Some obvious reasons are that (i) organic materials in inorganic ones reduce the overall density of the object, (ii) inorganic materials in organic ones improve their hardness and mechanical properties in elastomeric, glassy, and semicrystalline states, (iii) organic polymers in porous inorganics produce pore-free materials without high-temperature treatment, and (iv) mixing inorganic and organic materials on the nanoscale introduces new phenomena we are just beginning to understand.

In the following sections materials are described that are designed, in some cases, for structural applications and, in other cases, for functional uses, for optical, electronic, and catalytic applications. The processes used to prepare these materials variously lead to one-, two-, and three-dimensional structures. The phases are described as "guest and host," "infiltrant and matrix," and *in situ*. The common theme is that all of the materials are classified as hybrids and all of them are nanocomposites.

1. Definition of Nanocomposite

First, we use the term hybrid as the generic term for an inorganic–organic material. An accepted name for many of these materials is organically modified silicates (ormosils) (Schmidt and Wolter 1990) or, more generally, organically modified ceramics (ormocers) (Schmidt 1994). Another name is ceramic modified with polymer (ceramer) (Huang *et al.* 1985).

The second term we use commonly is nanocomposite to emphasize the nanometer level of mixing. Nanocomposites often refer to hybrids where organic and inorganic constituents are not covalently bonded, but it also applies to covalently bonded constituents. The forces acting between the constituents vary from weak to relatively strong.

Some notable attempts to classify hybrids have appeared. For example, Chujo and Saegusa (1992) have reviewed inorganic–organic materials prepared where the organic constituents are polymers, but not low molecular weight radicals or groupings. Sanchez and Ribot (1994) divide hybrids into two classes: Class I corresponds to hybrids where organic molecules are blended into the inorganic network, whereas Class II includes hybrids where inorganic and organic constituents are linked together via covalent bonds. Avnir *et al.* (1998) concentrate on inorganic–organic hybrids using organofunctional silicon alkoxides. Wen and Wilkes (1996) classify the resulting materials as networks. The dimensionality of hybrids has been reviewed by Lerner and Oriakhi (1997), principally in layered materials. We prefer the terms physical hybrids and chemical hybrids (Wojcik and Klein 1997). Overall these classifications organize hybrids into groups by looking at (i) the structure, (ii) methods of preparation, or (iii) the nature of links between organic and inorganic constituents.

The majority of processes treated in the following sections refer to the sol–gel process (Brinker and Scherer 1990). The sol–gel process is defined in this context as the hydrolytic process with a metal alkoxide such as tetraethylorthosilicate (TEOS). Upon addition of water TEOS in a cosolvent hydrolyzes:

$$\mathrm{Si(OC_2H_5)_4{+}4H_2O \rightarrow Si(OH)_4{+}4C_2H_5OH} \quad (3)$$

and polymerizes:

$$\mathrm{Si(OH)_4{+}Si(OH)_4 \rightarrow (OH)_3Si{-}O{-}Si(OH)_3{+}H_2O} \quad (4)$$

to build a silicate network. In time, the system arrives at the sol–gel transition where an irreversible gel exists. In the single component system, it is known that the result is an oxide skeleton which is porous once the solvent is removed.

The sol–gel approach as described is a room-temperature process. Polymers that otherwise are destroyed at ceramic processing temperatures can be incorporated. The impetus for so many combinations of polymers and oxides is, in fact, the sol–gel process. The concept of modifying inorganic networks by using covalent bonding works when the M–C (M is the metal species: aluminum, silicon, titanium, zirconium, etc.) bond is stable and does not undergo hydrolysis under the conditions of the sol–gel process.

1.1 Physical Nanocomposites

(*a*) *Interpenetrating networks.* A major classification of hybrids is interpenetrating networks, characterized by a time factor, i.e., sequential or simultaneous. The terminology of interpenetrating polymer networks (IPNs) is borrowed from organic polymer technology (Sperling 1981). In some cases semisequential is used to reflect the linear structure of the organic polymer in which a cross-linked inorganic network is formed.

Some of the earliest inorganic–organic hybrids were prepared by monomer infiltration into a previously formed silica gel. This is an example of a sequential interpenetrating network where the inorganic is the matrix. Specifically, when a dried silica gel (xerogel) is filled with monomer, the monomer can be polymerized *in situ*. Interpenetration is achieved when the impregnating monomer polymerizes in the open pores of the rigid, fully cross-linked silica matrix. In the polymer-impregnated gels, some porosity typically remains. Although some copolymers of methyl methacrylate, butadiene, and styrene have been used as polymers to impregnate silica, the best-known system is silica impregnated with PMMA (Klein 1994). While this type of hybrid was important at first, it has been surpassed by other methods that are simpler.

Another type of sequential interpenetrating network uses silica precursors (TEOS or tetramethylorthosilicate (TMOS)), which are able to solvate some organic polymers (Beaudry and Klein 1996). The polymer solubility enables the silica precursor to polymerize in the environment of an organic polymer solution. This is an example of a sequential interpenetrating network where the organic phase is the matrix. The number of polymers that form solutions with sol–gel formulations is, however, limited. Some initially soluble polymers tend to precipitate during gelation when a change in solvent composition leads to phase separation.

While only a few organic polymers are soluble in sol–gel formulations, many organic monomers are soluble in TEOS (Wojcik and Klein 1995a). These monomers can be introduced directly into sol–gel formulations. Both ring opening polymerizations and free radical polymerizations can be carried out simultaneously with the hydrolysis–condensation of TEOS. To reflect their synchronous polymerizations, the name is simultaneous inorganic–organic interpenetrating polymeric networks. This simultaneous route captures insoluble organic polymers within a sol–gel inorganic network. Generally, the resulting hybrids have no covalent linkages between the organic and inorganic components. The exception occurs when the organic monomer has hydroxyl or carboxyl groups that can lead to M–O–links.

In the preparation of simultaneous networks, it is important to control both polymerizations rates. On the one hand, systems with inorganic condensation rates much faster than the organic polymerization rates turn into brittle hybrids that shrink because the polymer content drops when unreacted monomer evaporates. On the other hand, systems with large organic polymerization rates usually show uncontrolled polymer precipitation. In practice, the kinetics of polymerization are difficult to control, so the success of the simultaneous approach rests on the careful selection of the monomers.

The idea of simultaneous interpenetrating networks has been advanced by Ellsworth and Novak (1993). Several monomers of acrylate type (free radical polymerization) or cyclic alkenes (ring opening metathesis) have been used. They include the acrylamide/*N*,*N*′-bisacrylamide monomer system, 2-hydroxyethyl or hydroxymethyl acrylate, and 7-oxanorbornene and its derivatives. In order to get nonshrinking hybrids, tetraacryloxyethoxysilane has been used instead of TEOS. The acrylate monomer liberated due to hydrolysis acts as a cosolvent to solvate the polymerizing silica network. The result is transparent hybrids with no shrinkage. Wojcik and Klein (1994) have studied the incorporation of silica with di- and triacrylate monomers, hexanediol diacrylate (HDDA), and glyceryl propoxy triacrylate (GPTA). These simultaneous IPNs are transparent hybrids with good mechanical strength.

For functional applications, inorganic gels can be doped with small quantities of organic molecules. Doped gels are prepared either by addition of organics into sol–gel solutions, or by infiltration of the dopant into previously formed xerogels. Although some interactions between the dopant molecules and the inorganic matrix take place, the optical properties or chemical functions of the organics are preserved in the solid state. A number of organics have been used, including organic dyes and chromophores, liquid crystals and active biomolecules, and enzymes and proteins, in order to produce specific optically active or bioactive devices for a broad range of optic and electro-optic applications. Suitably doped gels exhibit highly efficient luminescence, photochemical hole burning, photochromism, and tunable laser action (Avnir *et al.* 1994). In pure inorganic matrices, the activity of the dopant has been found to decrease because of thermodynamic incompatibility with the inorganic phase. Consequently, organic-doped inorganic gels have been replaced with organic-doped hybrid gels. A good example is PMMA–silica composites doped with organic laser dyes, where the hybrid has better properties as a laser host than either pure PMMA or pure silica.

(*b*) *Soluble polymers.* For those select polymers that are soluble in metal alkoxide solutions, the preparation of inorganic–organic hybrids is very simple. Taking again the example of silica, the polymerization of the silica precursor occurs around existing polymer chains or domains. Generally it is found that polymerization of TEOS in the presence of polymers under acidic conditions generates small, well-dispersed silica particles. Depending on the choice of polymer, highly transparent, lightweight materials can be obtained. Among the polymers used to form these sequential IPNs, those that give transparent hybrids are those capable of forming hydrogen bonds with hydroxylated silica particles (Landry *et al.* 1992a). Strong interactions between silanols, having the character of Brønsted acids, and specific groups on the polymer that are hydrogen acceptors are responsible for the high degree of two-phase mixing.

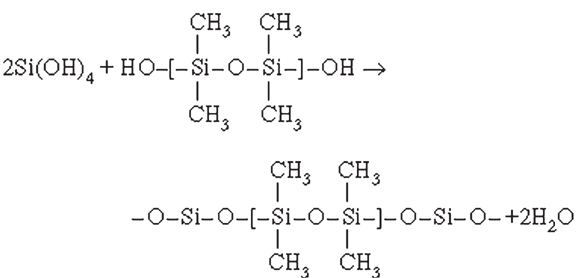
Linear polymers with hydrogen acceptor groups, amide, carbonyl, and carbinol, have been used in semisequential inorganic–organic IPNs. These polymers are poly(*N*,*N*-dimethacrylamide), poly(2-methyl-2-oxazoline), PMMA, poly(vinyl pyrrolidone) (PVP), poly(acrylic acid) (PAA) and its copolymers, PVAc, and poly(vinyl alcohol) (PVA) (Landry *et al.* 1992b).

In detail, the properties of the hybrids are functions of the relative fractions of the polymer and TEOS in the sol–gel formulation, along with the chemistry and length of the polymer chain. When TEOS is polymerized in the presence of the polymer, the glass transition temperature (T_g) of the polymer increases or becomes undetectable. The density of the hybrid does not follow a mixing rule for the inorganic and organic components, but is suppressed even with high inorganic contents. More or less porous structures can be formed with open or closed pores. We have studied the optical properties, mechanical properties, and phase separation in sol–gel formulations containing PVAc and poly(ethylene oxide) PEO. These are useful model systems for understanding the behavior in physical nanocomposites (Klein *et al.* 1997).

1.2 Chemical Nanocomposites

(*a*) *Monomer systems.* The purely inorganic three-dimensional network in the conventional sol–gel process consists of M–O–M units in the general case, or Si–O–Si units for silica. The purely organic polymer consists of C–C covalent bonds, in the general case, between mers in a polymer backbone, along with weaker van der Waals bonds between polymer chains. However, in a sol–gel process carried out with a precursor containing direct inorganic–organic links, it is possible to form homogeneous materials with C–C, M–C, and M–O bonds. This class of hybrids constitutes inorganic–organic polymers. Undoubtedly, the largest class of true inorganic–organic hybrids is made up of hybrids linked via Si–C bonds that are stable and do not undergo hydrolysis under the sol–gel processing conditions.

There are several methods to link covalently polymers to silica. In an early example, silanol-terminated poly(dimethylsiloxane) (PDMS) was co-condensed with TEOS:



$$2Si(OH)_4 + HO\text{–}[\text{–}\underset{CH_3}{\overset{CH_3}{Si}}\text{–}O\text{–}\underset{CH_3}{\overset{CH_3}{Si}}\text{–}]\text{–}OH \rightarrow -O-Si-O-[-\underset{CH_3}{\overset{CH_3}{Si}}-O-\underset{CH_3}{\overset{CH_3}{Si}}-]-O-Si-O- + 2H_2O \qquad (5)$$

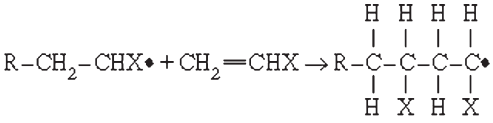
to give inorganic–organic PDMS chains, containing nonbridging methyl groups and nonlinear Si–O–Si angles, which exhibited significant flexibility (Mackenzie *et al.* 1992). When the functionality of PDMS is increased to six (the number of alkoxysilyl groups per oligomeric chain) or PDMS is replaced by poly(tetramethylene oxide) (PTMO), the stiffness and mechanical strength improve considerably (Wang *et al.* 1991).

In a way similar to PDMS and TEOS, other organic polymers have been functionalized with alkoxysilyl groups and covalently bonded to silica. Polymers that have been coupled with silica are PTMO-based polyurethane oligomers, polyoxazolines, polyimide (Mascia and Kioul 1995), poly(arylene ether ketone), poly(arylene ether sulfone), polystyrene, polyoxopropylene (PPO), polyacrylonitrile, copolymers of methyl methacrylate and allyl methacrylate (Wei *et al.* 1990), and cyclophosphazenes (Guglielmi *et al.* 1994).

Naturally, the properties of inorganic–organic hybrids are governed by their microstructure. For example, cellular structures are observed in PDMS–silica gels when the sol–gel parameters (water, HCl concentration, and temperature) are adjusted to accelerate the gelation process. With high concentrations of the organic component, hybrids have low densities, indicating high free volumes characteristic of the organic polymers.

Another method leading to covalent linkage of organic polymer to silica is free radical polymerization. When an organic radical R′ in an organo-substituted silicoester of formula $R'Si(OR)_3$ contains olefinic moieties, it is possible to build a polymeric network by thermal or photochemical free radical polymerization of C=C bonds, for example:

$$R\bullet + CH_2 = CHX \rightarrow R-CH_2-CHX\bullet \qquad (6)$$



$$R-CH_2-CHX\bullet + CH_2{=}CHX \rightarrow R-\underset{H}{\overset{H}{C}}-\underset{X}{\overset{H}{C}}-\underset{H}{\overset{H}{C}}-\underset{X}{\overset{H}{C}}\bullet \qquad (6a)$$

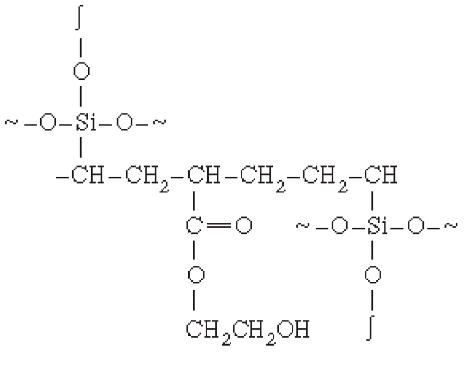
where R• is a free radical and X is Cl, acetate, CH_2CCl_2, etc. An initiator to produce the free radical is often benzoyl peroxide, $(C_6H_5)_2C_2O_4$. Epoxide groups are also polymer-formers via cationic polymerization or via condensation with curing agents bearing groups such as amine or hydroxyl. Organic polymerization can be carried out simultaneously with condensation of hydrolyzed alkoxysilyl groups.

The commonly used organoalkoxysilanes bearing polymerizable C=C bonds are 3-methacryloxypropyl trimethoxysilane (MEMO) ($CH_2{=}C(CH_3)COO(CH_2)Si(OCH_3)_3$) and vinyl trimethoxysilane (VTMS) ($CH_2{=}CH\text{–}Si(OCH_3)_3$) or their ethoxy analog. Alkoxysilyl groups of these inorganic–organic monomers form Si–O–Si bonds upon condensation or co-condensation with another silica precursor such as TEOS. At the same time, (meth)acryl or vinyl groups are able to polymerize with inorganic–organic monomers or with (meth)acrylic or vinyl monomers. These processes lead to the formation of inorganic–organic polymers and copolymers.

Reportedly, inorganic polycondensations of inorganic–organic monomers by the sol–gel process from single monomer systems are difficult. For most of the inorganic–organic precursors, gelation times under acidic conditions are quite long. The reason for the long gelation time is the low concentration of alkoxysilyl groups and steric hindrance from organics in the system. However, the process can be accelerated by incorporation of TEOS.

An idealized structure of an inorganic–organic copolymer synthesis by co-condensation and copolymerization of the tri-monomer system TEOS–vinyltriethoxysilane (VTES)–hydroxyethyl methacrylate

(HEMA) is shown below. To accelerate gelation and at the same time reduce shrinkage, inorganic oligomers are used instead of inorganic precursors.



(7)

Ellsworth and Novak (1993) used soluble oligomers of silicic acid instead of TEOS in a condensation with (meth)acryl trialkoxysilane to increase the silica content of inorganic–organic hybrids and reduce shrinkage. Hoebbel *et al.* (1994) used tetramethylammonium silicate ($(N(CH_3)_4)_8Si_8O_{20}\bullet 69H_2O$) converted by an ion-exchange process into double four-ring oligomers of silicic acid ($H_8Si_8O_{20}$). This oligomer can be stabilized by reaction with functional siloxanes containing polymerizable groups like vinyl, allyl, hydrido, or chloromethyl. The functionalized derivative can be polymerized while preserving silicic acid cages in the structure.

When the inorganic–organic monomer bears two or more olefinic moieties in its structure, the monomer becomes highly reactive, creating very favorable conditions for rapid cross-linking by photopolymerization. A series of diacrylate alkoxysilanes was synthesized by Wolter *et al.* (1992) by coupling mercapto-substituted alkoxysilanes to the triacrylate monomers or acrylate resins. The resulting inorganic–organic monomer can be co-condensed via its alkoxysilyl groups with TEOS, whereas its acrylic groups can be cured by UV light, or polymerized thermally. Likewise, diacrylotrialkoxy silane inorganic–organic monomer from aliphatic triacrylate-glycerol propoxy triacrylate (GPTA) and 3-mercaptopropyl trimethoxysilane was used to synthesize UV-curable coatings for optical fibers (Wojcik *et al.* 1993).

A subgroup of hybrids containing a covalent link are those where the polymer exhibits specific electronic or electro-optic properties. Examples of these hybrids include nonlinear optics (NLO) polymer–silica gels, organic dye–silica gels, and conducting polymer–silica gels. For example, organic dye–silica gels of cyan, magenta, and yellow were synthesized at Kodak, where the chromophore was linked covalently to the polymeric component via the reaction of an amine, halogen, or isocyanate functionality, while the silica and polymeric components were covalently bonded via the action of 3-chloropropyltrimethoxysilane. Polymers with conducting properties like polypyrrole, polyaniline, or polythiophene have been covalently incorporated into silica, following their alkoxysilyl functionalization (Sanchez *et al.* 1994).

(*b*) *Monomer–polymer systems*. When the organic component is a low molecular weight group or radical and not a polymeric chain or network, the hybrid is an organically modified gel (ormogel) as distinct from an ormosil. Ormogels are formed from $R'Si(OR)_3$-type organosilanes where radical R has a network-modifying effect and is used to introduce new functionalities into the inorganic network. For example, hydrophobicity can be achieved by introducing a methyl, phenyl, or fluoroalkyl chain for R (Schmidt *et al.* 1986). Polymerization of $R'Si(OR)_3$ alone leads to formation of silsesquioxanes of general formula $(RSiO)_{1.5}$, a silicate framework where each silicon is linked covalently to an organic radical R. A new class of hybrids is emerging called polyhedral oligomeric silsesquioxanes (POSS) that can be blended into rubbers, resins, thermoplastics, and thermosets to enhance thermal stability (Lichtenhan *et al.* 1995).

Bifunctional silanes of formula $R'R''Si(OR)_2$ are used also to prepare ormogels. Upon polymerization, these organoalkoxysilanes form cyclic species or linear chains that are not able to cross-link and form networks themselves. They can be used with tetrafunctional silanes (TEOS) to incorporate desirable functionality into the hybrid network. There are also ormogels where radical R forms a bridge between two silicon atoms. Good examples of these are arylene or alkylene-bridged polysilsesquioxanes synthesized from bis(triethoxysilyl) or bis(trichlorosilyl) aryl or alkyl monomers (Shea *et al.* 1989). Various aliphatic or aromatic molecules have been used as the organic spacer between two silicon atoms.

(*c*) *Copolymers*. Copolymers are the products of addition polymerization between two different monomers. Some examples have been mentioned in the sections on physical nanocomposites, as well as chemical nanocomposites. Between two linear polymers such as PVC and PVAc, the units can be arranged in a random copolymer, block copolymer, or graft copolymer. The interesting aspect of copolymers is that the properties of the copolymer are often quite different from the properties of either homopolymer.

In the context of nanocomposites, the term copolymer is applied to polymers with organic and inorganic units. Here especially, the properties of the copolymer are different from the properties of either component. Systems with two or more monomers open opportunities for designing properties, using surface treatments, coupling agents, and sophisticated manufacturing methods.

To recap on chemical nanocomposites, we find that covalently bonded hybrids combine the flexibility and mechanical properties of the organic polymer chains with the hardness and stiffness of the inorganic polymer. In detail, the properties of the hybrids are governed by the chemistry, presence of side chains or pendant groups, and length of polymeric chains. Especially important are the concentration and distribution of alkoxysilyl or equivalent groups, as they directly affect the rigidity or flexibility of the system.

2. Classes of Polymers in Nanocomposites

The five classes of polymers that have been incorporated into organic–inorganic hybrids, principally using the sol–gel process, are discussed.

2.1 Elastomers

An elastomer is a polymer with high elasticity due to its coiled structure. A typical elastomer is nonvulcanized rubber, where the chains are first unkinked when a stress is applied, followed by chain straightening and lengthening. Elastomers are poorly crystallized but become more crystalline as they are stretched. Some examples of elastomers include rubbers, polyurethane, and PDMS.

2.2 Partially Crystalline Polymers

Crystallization in polymers is often difficult, leading to glassy or partially crystalline polymers. In general, linear polymers are easier to crystallize than network polymers. Isotactic (the same stereostructure) polymers are easier to crystallize than atactic (random positions) or syndiotactic (alternating positions) polymers. Likewise, it is easier to crystallize polymers in the *trans* (unsaturated positions on opposite sides of the chain) than in the *cis* (unsaturated positions on same side of the chain) configuration. Condensation polymers such as poly(ethylene glycol terephthalate) (Dacron), polyimide, and nylon exhibit some crystallinity because of their moderately stiff chain structures.

2.3 Glassy Polymers

A common example of a glassy polymer is PMMA, where it is difficult to crystallize PMMA because of the atactic structure following addition polymerization. Glassy polymers exhibit a glass transition characterized by the temperature range where the polymer changes its behavior from glassy to rubbery.

2.4 Liquid-Crystalline Polymers

An example of a liquid-crystalline (LC) polymer is hydroxypropyl cellulose (HPC) or 4′-pentyl-4-biphenylcarbonitrile. The characteristic of LC polymers is that they align what are called nematogenic compounds in response to electric fields. LC polymers encapsulated in inorganic materials have been developed for displays. The inorganic encapsulation gives the liquid crystals thermal stability.

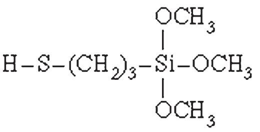
2.5 Thermosets

Thermosetting polymers are polymers that do not soften when heated. Instead, a thermoset polymerizes further with increasing temperature, resulting in a continuous network where polymer units cannot slip past each other. An example of a thermosetting polymer is phenol formaldehyde, where phenol (C_6H_6O) and formaldehyde (CH_2O) undergo condensation polymerization to generate phenol formaldehyde that is able to link at a minimum of three places and a maximum of five places.

3. Weak and Strong Bonding in Nanocomposites

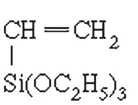
3.1 Covalently Bonded Systems

The majority of hybrids investigated have covalent bonds between the organic and inorganic components. These bonds are formed before incorporation into a nanocomposite. A typical route involves a metal alkoxide that has been functionalized via a mercapto group (Wojcik and Klein 1995b):



$$H-S-(CH_2)_3-\underset{\displaystyle OCH_3}{\overset{\displaystyle OCH_3}{\overset{|}{\underset{|}{Si}}}}-OCH_3 \qquad (8)$$

which is used during the creation of the nanocomposite when a reactive site, such as a vinyl group:



$$\begin{array}{l} CH{=}CH_2 \\ | \\ Si(OC_2H_5)_3 \end{array} \qquad (9)$$

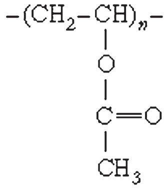
is available and covalently bonded to the silicon atom. The reason that forming direct linkages has been popular is, in large part, that it ensures strong bonds between the components.

3.2 Hydrogen-Bonded Systems

The minority of hybrids investigated have hydrogen bonds between the organic and inorganic components. These bonds are formed between hydroxyls on the inorganic gel and ether oxygens on the polymers assuming a resonant hybrid pair for PEO:

$$[+CH_2-CH_2-O-/+O-CH_2-CH_2-]\rightarrow -(O-CH_2-CH_2-)_{n}- \qquad (10)$$

or carbonyl groups on the polymers:



$$-(CH_2-\underset{\underset{\underset{\underset{CH_3}{|}}{C=O}}{\underset{|}{O}}}{\underset{|}{CH}})_n- \qquad (11)$$

as in the case of PVAc. The reason that there are few hydrogen-bonded hybrids is directly related to the small number of water-soluble monomers and the even smaller number of soluble polymers. Nevertheless, the technological importance of the hydrogen-bonded systems is evident in the development of organically modified electrolytes (ormolytes) (Charbouillot *et al.* 1988).

3.3 Phase Separation: Phenomena and Kinetics

As described above, physical hybrids depend on the polycondensation of an inorganic network in the presence of an organic polymer dissolved in the solution, or polymerization of organic monomers inside the porous inorganic network. To obtain homogeneous mixing, the degree of phase separation has to be controlled. There should be strong, attractive interactions between the organic and inorganic polymers to achieve good interfacial bonding. Polymers containing carbonyl groups, such as PVAc, can be incorporated into sol–gel processes because they are soluble in alcohol–water mixtures due to partial esterification of the side chains. Polymers containing ether oxygens, such as PEO, sometimes referred to as poly(ethylene glycol) (PEG), are soluble in water because the ether oxygens readily form hydrogen bonds.

Landry *et al.* (1992b) have demonstrated the simplicity of the approach with TEOS and a number of polymers. Depending on (a) the synthesis conditions and (b) the relative amounts of organic and inorganic polymer, the morphologies ranged from discrete silica particles within an organic matrix, to interpenetrating networks, to polymeric domains dispersed within a silica matrix. Under acidic conditions with TEOS it was possible to form optically transparent films with PMMA, PVAc, poly(vinyl pyrrolidone), and poly(*N*,*N*-dimethacrylamide). Hydrogen bonding between the silanols on the silica polymer and the carbonyl groups on the polymers prevented macroscopic phase separation. Phase separation, if any, occurred on such a small scale that transparent materials were formed without the need for covalent bonding. Furthermore, there is no evidence for chemical reactions below 150 °C. Acetic acid was evolved from PVAc above 200 °C, but this is likely to be a decomposition product and not an indication of covalent bonding between TEOS and PVAc.

Polymerization of TEOS in the presence of a soluble polymer does not always lead to a homogeneous mixture. Kaji *et al.* (1993) have taken advantage of the incompatibility of PAA and silica and PEO and silica. Their studies have shown that the scale of the phase separation and its timing are dependent on (i) the amount of polymer, (ii) the molecular weight of the polymer, (iii) the solvent composition, and (iv) the pH. The gel morphologies are the results of a competition between phase separation and gelation, also referred to as "chemical quenching." The rate of gelation of the silica network is in opposition to the rate of phase separation, which is an indication of the segregation strength. When gelation is faster than phase separation, the system becomes rigid with a morphology reflecting the early structure of the gel. When phase separation is faster than gelation, the system becomes rigid with a variety of transient structures that are increasingly heterogeneous, with the extreme case being sedimentation.

4. Importance of Interfaces in Inorganic–Organic Nanocomposites

The interface between the organic and the inorganic constituents is the key issue in hybrids and nanocomposites. In our work we are measuring strengths, hardness, and skeletal densities of some representative systems. Mechanical behavior is at best an indirect measurement of interface behavior. However, by relating these properties to bonding and microstructure, we try to examine the nature of the interface and the scale on which it controls properties.

Other properties, such as optical, electronic, magnetic, catalytic, and chemical durability require further investigation. We need to understand what these properties tell us about the interfaces, so that we can better engineer the interfaces.

Finally, we should address the questions, what analytical techniques are available to probe the interface? Are indirect measurements more or less meaningful than direct measurements, considering that performance may be a reflection of averaged properties? Other articles in this encyclopedia should shed some light on these subjects.

Bibliography

Andrews M P, Najafi S I (eds.) 1997 *Critical Reviews: Sol–Gel and Polymer Photonic Devices* (CR68-03). SPIE, Bellingham, WA

Avnir D, Braun S, Lev O, Levy D, Ottolenghi M 1994 Organically doped sol–gel porous glasses: chemical sensors, enzymatic sensors, electrooptical materials, luminescent materials and photochromic materials. In: Klein L C (ed.) *Sol–Gel Optics*. Kluwer, Boston, pp. 539–82

Avnir D, Klein L C, Levy D, Schubert U, Wojcik A B 1998 Organo-silica sol–gel materials. In: Rappoport Z, Apeloig Y (eds.) *The Chemistry of Organosilicon Compounds*. Wiley, London, Vol. 2, pp. 2317–62

Beaudry C L, Klein L C 1996 Sol–gel processing of silica/poly(vinyl acetate) (PVAc) nanocomposites. In: Chow G-M, Gonsalves K E (eds.) *Nanotechnology: Molecularly Designed Materials*. American Chemical Society, Washington, DC, pp. 382–94

Brinker C J, Scherer G W 1990 *Sol–Gel Science*. Academic Press, Boston

Charbouillot Y, Ravaine D, Armand M, Poinsignon C 1988 Aminosils: new solid state protonic materials by the sol–gel process. *J. Non-Cryst. Solids* **103**, 325–30

Chujo Y Saegusa T 1992 Organic polymer hybrids with silica gel by means of the sol–gel method. In: *Advances in Polymer Science*. Springer, Berlin, Vol. 100, pp. 12–29

Coltrain B K, Sanchez C, Schaefer D W, Wilkes G L (eds.) 1996 *Symp. Proc. 435, Better Ceramics Through Chemistry VII: Organic/Inorganic Hybrid Materials*. Materials Research Society, Pittsburgh, PA

Drzal L T, Opila R L, Peppas N A, Schutte C (eds.) 1995 *Symp. Proc. 385, Polymer/Inorganic Interfaces II*. Materials Research Society, Pittsburgh, PA

Ellsworth M W, Novak B M 1993 "Inverse" organic–inorganic composite materials: 3. High glass content "nonshrinking" sol–gel composites via poly(silicic acid esters). *Chem. Mater.* **5**, 839–44

Guglielmi M, Colombo P, Brusetin G, Facchin G, Gleria M 1994 New materials based on the reaction of cyclo- and poly-(organo phosphazenes) with SiO_2, TiO_2 and ZrO_2 precursors. *J. Sol–Gel Sci. Technol.* **2**, 109–14

Hoebbel D, Endres K, Reinert T, Pitsch I 1994 Inorganic–organic polymers derived from functional silicic acid derivatives by additive reaction. *J. Non-Cryst. Solids.* **176**, 179–85

Huang H H, Orler B, Wilkes G L 1985 Ceramers: hybrid materials incorporating polymeric/oligomeric species with inorganic glasses by a sol–gel process: 2. Effect of acid content on the final properties. *Polym. Bull.* **14**, 557–64

J. Sol–Gel Sci. Technol. 1995 **5**(2), 75–157

J. Sol–Gel Sci. Technol. 1996 **7**(3), 151–224

Kaji H, Nakanishi K, Soga N 1993 Polymerization-induced phase separation in silica sol–gel systems containing formamide. *J. Sol–Gel Sci. Technol.* **1**, 35–46

Klein L C 1994 Nanocomposite fabrication for transparent windows. In: Klein L C (ed.) *Sol–Gel Optics*. Kluwer, Boston, pp. 215–32

Klein L C, Beaudry C L, Yamazaki S, Wojcik A B 1997 Transparent polymer/silica hybrid gels. In: Andrews M P, Najafi S I (eds.) *Sol–Gel and Polymer Photonic Devices*. SPIE, Bellingham, WA, pp. 54–63

Klein L C, Francis L F, DeGuire M R, Mark J E (eds.) 1999 *Symp. Proc. 576, Organic/Inorganic Hybrid Materials II*. Materials Research Society, Pittsburgh, PA

Laine R M, Sanchez C, Brinker C J, Giannelis E (eds.) 1998 *Symp. Proc. 519, Organic/Inorganic Hybrid Materials*. Materials Research Society, Pittsburgh, PA

Landry C J T, Coltrain B K, Brady B K 1992b In situ polymerization of tetraethoxysilane in poly(methyl methacrylate). *Polymer* **33**, 1486–95

Landry C J T, Coltrain B K, Wesson J A, Zumbulyadis N, Lippert J C 1992a In situ polymerization of tetraethoxysilane in polymers. *Polymer* **33**, 1496–506

Lerner M M, Oriakhi C O 1997 Polymers in ordered nanocomposites. In: Goldstein A N (ed.) *Handbook of Nanophase Materials*. Marcel Dekker, New York, pp. 199–219

Lichtenhan J D, Otonari Y A, Carr M J 1995 Linear hybrid polymer building blocks: methacrylate-functionalized polyhedral oligomeric silsesquioxane monomers and polymers. *Macromolecules* **28**, 8435–7

Mackenzie J D, Chung Y J, Hu Y 1992 Rubbery ormosils and their applications. *J. Non-Cryst. Solids* **147/148**, 271–9

Mascia L, Kioul A 1995 Influence of siloxane composition and morphology on properties of polyimide–silica hybrids. *Polymer* **19**, 3649–59

Sanchez C, Alonso B, Chapusot F, Ribot F, Audebert P 1994 Molecular design of hybrid organic–inorganic materials with electronic properties. *J. Sol–Gel Sci. Technol.* **2**, 161–6

Sanchez C, Ribot F 1994 Design of hybrid organic–inorganic materials synthesized via sol–gel chemistry. *New J. Chem.* **18**, 1007–47

Schmidt H 1994 Inorganic–organic composites by sol–gel techniques. *J. Sol–Gel Sci. Technol.* **1**, 217–31

Schmidt H, Scholze H, Tunker G 1986 Hot melt adhesives for glass containers by the sol–gel process. *J. Non-Cryst. Solids* **80**, 557–63

Schmidt H, Wolter H 1990 Organically modified ceramics and their applications. *J. Non-Cryst. Solids* **121**, 428–35

Shea K J, Loy D A, Webster O W 1989 Aryl-bridged polysilsesquioxanes—new microporous materials. *Chem. Mater.* **1**, 572–4

Sperling L H 1981 *Interpenetrating Polymer Networks and Related Materials*. Plenum, New York

Wang B, Wilkes G L, Hedrick J C, Liptak S C, McGrath J E 1991 New high refractive index organic/inorganic hybrid materials from sol–gel processing. *Macromolecules* **24**, 3449–53

Wei Y, Bakthavatchalam R, Whitecar C K 1990 Synthesis of new organic–inorganic hybrid glasses. *Chem. Mater.* **2**, 337–9

Wen J, Wilkes G L 1996 Organic/inorganic hybrid network materials by the sol–gel approach. *Chem. Mater.* **8**, 1667–81

Wojcik A B, Klein L C 1994 Organic–inorganic gels based on silica and multifunctional acrylates. *J. Sol–Gel Sci. Technol.* **2**, 115–20

Wojcik A B, Klein L C 1995a Transparent inorganic/organic copolymers by the sol–gel process: thermal behavior of copolymers of tetraethyl orthosilicate (TEOS), vinyl triethoxysilane (VTES) and (meth) acrylate monomers. *J. Sol–Gel Sci. Technol.* **5**, 77–82

Wojcik A B, Klein L C 1995b Transparent inorganic/organic copolymers by the sol–gel process: copolymers of tetraethyl orthosilicate (TEOS), vinyl triethoxysilane (VTES) and (meth)acrylate monomers. *J. Sol–Gel Sci. Technol.* **4**, 57–66

Wojcik A B, Klein L C 1997 Organic/inorganic hybrids by the sol–gel process: a classification scheme. *Appl. Organomet. Chem.* **11**, 129–35

Wojcik A B, Klein L C, Rondinella V V, Matthewson M J, Foy P 1993 Organically modified silicate coatings for optical fibers. SPIE 2074, pp. 135–41

Wolter H, Glaubitt W, Rose K 1992 Multifunctional (meth)acrylate alkoxysilanes: a new type of reactive compounds. In: Hampden-Smith M J, Klemperer W G, Brinker C J (eds.) *Symp. Proc. 271, Better Ceramics Through Chemistry V*. Materials Research Society, Pittsburgh, PA, pp. 719–24

L. C. Klein[a] and A. B. Wojcik[b]
[a] *Rutgers University, Piscataway, New Jersey, USA*
[b] *Hybrid Glass Technologies, Princeton, New Jersey, USA*

Polymer-modified Ceramics

Ceramics are inorganic solid compounds that contain metallic and nonmetallic elements, such as silica glass, the cement of concrete, and most of high temperature superconductors. Polymers, such as polystyrene, polysiloxane, and proteins, are large molecules built up by the linking together of great numbers of much smaller molecules that contain nonmetallic elements. Polymer-modified ceramics are materials that contain both the polymer and ceramic components mixed homogeneously at the molecular or nanometer scales. There are two major families of polymer-modified ceramics: (i) polymer–ceramic hybrid materials, which have strong chemical interactions between the polymer and ceramic components such as extensive hydrogen-bonds and/or covalent bonds; and (ii) polymer–ceramic molecular composites (or nanocomposites), in which the polymer components are physically trapped in the ceramic networks with no or limited chemical interactions. Though all these terms (and many others) have been loosely used, one significant feature of the polymer-modified ceramics is that the polymer components are uniformly dispersed, with or without formation of chemical bonding, in the inorganic networks and, therefore, the macroscopic interfaces (e.g., $\geqslant 400$ nm—wavelength of visible light), that exist in traditional composite materials between the polymer and the inorganic components, are eliminated.

In this article, the major theme is the synthesis and characterization of polymer-modified ceramic materials in which the polymer is the dispersed phase for improving the properties of the host ceramic matrix. Discussion is focused mainly on the hybrid materials that have strong interactions, such as extensive hydrogen-bonding and/or covalent bonding, between the polymer (particularly vinyl plastics) and ceramic (particularly silica, alumina, and titania) components. Various synthetic routes are reviewed with emphasis on the chemically or photochemically catalyzed sol–gel reactions under controlled conditions, so that the polymer chains are uniformly incorporated into the ceramic networks without macroscopic phase separation. The chemical, physical and biological properties of the polymer-modified ceramics are described and their dependence on the structure and composition of the polymer and ceramic components is outlined. Examples are selected mostly from results, including some original data, from the author's laboratory. A number of potential applications are presented. Related information can be found in *Polymer–Ceramic Nanocomposites: Polymer Overview** and *Polymer–Ceramic Nanocomposites: Ceramic Phases**.

1. General Principles

Since the 1970s, the field of materials science has witnessed the emergence of a sol–gel process for the preparation of ceramic materials from precursor compounds in the presence of a catalyst at relatively low temperatures. For example, acid-catalyzed hydrolysis and polycondensation (i.e., the sol–gel reactions) of tetraethyl orthosilicate (TEOS) yield a silica glass at temperatures much lower than the conventional fusion process. Because most organic polymers (i.e., carbon-containing polymers) undergo thermal decomposition under the high-temperature conditions that are often required for processing ceramics, the sol–gel process opened a new door in materials science that allows for the synthesis of polymer-modified ceramics, i.e., the integration of polymers and ceramics, two major types of industrial materials in modern society.

From 1980 to 1990, silica sol–gel materials were covalently modified by small organic groups, i.e., organic-modified silica or Ormosils (e.g., Schmidt 1985), by polymers containing heteroatoms in the backbones such as polydimethylsiloxanes (e.g., Huang *et al.* 1985), polyethers (e.g., Huang and Wilkes 1987), and polyoxazolines (e.g., Chujo *et al.* 1990), and by vinyl polymers such as polymethacrylates (e.g., Wei *et al.* 1990). Modification of polymers by *in situ* precipitation of sol–gel silica was also achieved (e.g., Jang and Mark 1984). To date, many dozens of various types of polymers have been successfully incorporated into silica, titania, alumina, zirconia and other ceramic matrices via the sol–gel process. A number of excellent books and reviews have been published related to the polymer-modified ceramics (e.g., Mackenzie and Ulrich 1988, Brinker and Scherer 1990, Wilkes *et al.* 1990, Novak 1993, Mark *et al.* 1995, Laine *et al.* 1998, Klein *et al.* 1999).

Ceramics are generally known for their hardness and brittleness, along with their resistance to high temperatures and to severe environments. In addition, many ceramic materials such as silica glass have excellent optical properties such as transparency. For most applications, their brittleness (i.e., lack of impact strength) is the major, sometimes fatal, deficiency of ceramics. However, organic polymers are usually noted for their low density and high toughness (i.e., high impact strength). They can be tailor-made to have excellent elasticity (e.g., synthetic rubbers) or optical transparency (e.g., polymethacrylates or Plexiglas$^{\mathrm{TM}}$). However, lack of hardness is one of most significant flaws of polymers in many applications. Associated with the lack of hardness are the problems of low wear and scratch resistance as well as dimensional stability. The development of conventional composite materials with ceramics as fillers and polymers as matrices have solved some of the above-mentioned problems to some extent. Because of organic–inorganic thermodynamic incompatibility, the presence of macroscopic interfaces between the inorganic fillers and organic matrices often leads to many undesired problems in the performance of materials. For example, mechanical failure often occurs because of cracking of the interfaces. Treatment of the fillers with silane coupling agents strengthens the interfaces to a certain but limited extent. In addition, optical properties such as transparency of either silica or polymethacrylates cannot be maintained due to the existence of macroscopic phase separation in conventional composites.

Hence, among the motivations for developing polymer-modified ceramics has been toughening of ceramics by polymers or hardening of polymers by ceramics in a controllable manner without sacrificing the optical properties. By varying the structure and composition of polymer and ceramic components, the idea was that properties of the polymer-modified ceramics would be designable in a continuum between those of pure polymers and pure ceramics. Polymer–ceramic nanocomposites or molecular composites would achieve such a goal. If the polymer–ceramic phase separation is at the microscopic level (e.g., domain size of either the polymer or ceramic component $<400\,\mathrm{nm}$), the materials are transparent to visible light. However, again because of the thermodynamic incompatibility between the usually hydrophobic polymer chains and the usually hydrophilic ceramic matrices, the dimension of the phase separation cannot be controlled easily during the sol–gel processing and the interfaces, however small in scale, remain weak. In contrast, polymer–ceramic hybrid materials could fundamentally eliminate the interface problem because the thermodynamic repulsion can be compensated by the formation of strong chemical bonding between the polymer and ceramic components. In the case where the bonding is covalent, the materials can be considered as cross-linked copolymers that contain organic and inorganic building blocks.

2. *Synthesis*

There are a number of general strategies for the synthesis of polymer-modified ceramics based on sol–gel chemistry. (i) Ceramic matrix route: forming a porous inorganic sol–gel matrix first and impregnating the matrix with monomers followed by polymerization *in situ*. (ii) Polymer matrix route: forming a polymer matrix first and impregnating it with an inorganic precursor followed by the sol–gel reactions to deposit or precipitate a ceramic component in the polymer matrix. (iii) Polymer precursor route: forming a polymer precursor that bears sol–gel active functional groups first and mixing it with an inorganic precursor followed by the sol–gel reactions of both precursors. (iv) Simultaneous formation of ceramics and polymers: preparing a mixture or compound that contains both the sol–gel reactive preceramic and the polymerizable monomer components first and initiating both the sol–gel reactions and polymerization at nearly the same time.

It is noted that the ceramic materials prepared via the sol–gel process without densification at elevated temperatures are porous and, hence, often referred to as gels. For example, pure silica obtained by the sol–gel reactions of an inorganic precursor like TEOS under acid catalysis at ambient temperature is microporous, with internal pore diameter and surface area typically about 15–20 Å or smaller and about $100–200\,\mathrm{m^2\,g^{-1}}$ respectively, based on measurements of adsorption isotherms of gases such as N_2. The ceramic matrix route takes advantage of this porosity by impregnating monomers along with polymerization initiators or catalysts into the pores. Polymerization of the monomers inside the porous ceramic matrix yields a polymer-modified ceramic material, which can be considered as a polymer-ceramic semi-interpenetrating network or interpenetrating network (IPN) if the polymerization results in cross-linking of polymer chains. As an example, monolithic sol–gel silica was soaked with methyl methacrylate (MMA), which was then polymerized *in situ* with a free radical initiation system to afford the PMMA-modified silica (Pope *et al.* 1989). It is conceivable that further exploration of this route would lead to many interesting new materials, particularly if some polymerizable functional groups are present on the surface of the internal pores of ceramic matrices to form covalent bonds between the matrices and the polymer chains, and if the pore parameters, e.g., size and distribution of the internal pores in the matrices, can be controlled as in the cases of microporous and mesoporous molecular sieves.

In the polymer matrix route, a polymer with or without cross-linking is impregnated with an inorganic sol–gel precursor, which undergoes the sol–gel reactions *in situ* to form ceramic components in the polymer matrix. As examples, sol–gel silica was successfully introduced into plastics and elastomers (Jang and Mark 1984). As a possible generalization of this polymer matrix route, a polymer could be mixed with a sol–gel precursor to form a homogeneous solution and, after the sol–gel reactions, the polymer chains could be entrapped in the ceramic network, in principle, uniformly at the molecular level. However, among the difficulties in such a generalization is the thermodynamic incompatibility of organic polymers with inorganic networks that leads to macroscopic phase separation before and/or after the gelation (i.e., in sols and in gels, respectively), as often evidenced by the formation of translucent or opaque gel products. There are a few examples where highly hydrophilic polymers such as polyethers are encapsulated in inorganic sol–gel matrices without apparent macroscopic phase separation. To prevent the macroscopic phase separation, strong interactions between the polymer and ceramic components are required—this leads to the third and the fourth synthetic strategies, i.e., the polymer precursor route and the simultaneous formation of polymers and ceramics, which are described below in detail.

2.1 Polymer Precursor Route

In the polymer precursor route, preformed polymers that contain a designed number of sol–gel active functionalities are employed as precursors to mix and react with inorganic sol–gel precursors. The active functionalities include those that can undergo the sol–gel reactions with the inorganic precursors, such as hydroxysilyl (≡Si-OH), alkoxysilyl (≡Si-OR), silane (≡Si-H), and halosilyl (≡Si-X; X = Cl, Br, etc.) groups. The cross-condensation of these functional groups with similar or identical groups in the inorganic precursor, e.g., $Si(OEt)_4$ and its hydrolysis product $Si(OH)_4$, would result in covalent bonding, e.g., siloxane ≡Si-O-Si≡bonds, between the polymer chains and the ceramic matrix, e.g., silicates formed by self-condensation of $Si(OEt)_4$ and/or $Si(OH)_4$. Since the formation of covalent bonds may kinetically pre-empt the thermodynamic repulsion, hydrophobic polymers like vinyl polymers could be homogeneously transfixed in hydrophilic ceramic matrices. A general methodology has been thus developed to prepare vinyl polymer-modified ceramic hybrid materials (Wei *et al.* 1990, 1995a).

The use of vinyl polymers to modify ceramics is of great importance because these polymers are among the most widely used organic materials in our daily life. To list just a few general categories of usage, polystyrene (PSt), poly(methyl methacrylate) (PMMA), poly(vinyl chloride) (PVC), and isotactic polypropylene can be used as plastics; polyacrylonitrile (PAn), poly(vinyl acetate) (PVAc), and linear polyethylene as fibers; and polybutadiene, polyisoprene, and poly(styrene-block-butadiene) (SBR) as elastomers or rubbers. As a basic strategy for the synthesis of the sol–gel active vinyl polymers such as poly[methyl methacrylate-co-3-(trimethoxysilyl)propyl methacrylate] [P(MMA-MSMA)], common vinyl monomers such as methyl methacrylates, acrylonitrile, and styrene are copolymerized with monomers such as 3-(trimethoxysilyl)propyl methacrylate (MSMA) that contain one or more sol–gel active functional groups, e.g., -SiOMe. The general equations for the preparation of vinyl polymer precursors and their subsequent sol–gel reactions with inorganic precursors are illustrated in Fig. 1.

The number (y) of the sol–gel active units in the polymer precursor can be designed and controlled by varying the co-monomer feed and other synthetic parameters. It is usually maintained low to yield the copolymers whose properties are closer to those of conventional vinyl polymers. However, if it is too low, macroscopic phase separation might occur during the sol–gel reactions. The distribution of the sol–gel active units along the copolymer chains can also be controlled by using proper copolymerization

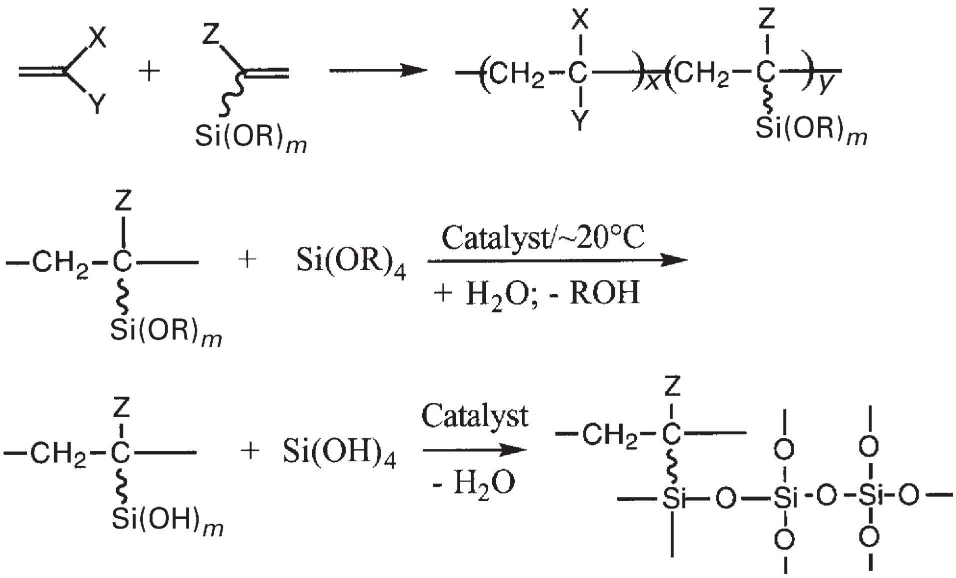


Figure 1
Preparation of the sol–gel active polymer precursors and subsequent sol–gel reactions with inorganic precursors. X, Y and Z are various groups; R: alkyl groups; $m = 1$, 2 or 3 (with attachment of other groups to Si when $m \neq 3$); Si can be replaced by other metals such as Al, Ti, Zr, etc.; -OR can be replaced by other sol–gel active groups such as halogen, etc.

methods and conditions. These polymer precursors are then hydrolyzed and co-condensed with one or more inorganic precursors at ambient temperature to afford the vinyl polymer-modified hybrid ceramic materials. The generality of this strategy could in principle allow for incorporation of virtually any vinyl polymers into sol–gel ceramic matrices.

The alkoxysilyl-containing monomers that have been employed include 3-(trimethoxysilyl)propyl methacrylate, 3-(triethoxysilyl)propyl methacrylate (ESMA), styrylethyltrimethoxysilane (STMS), etc. It is interesting to note that many of these sol–gel active monomers have been extensively used as silane coupling agents in the manufacturing of traditional composite materials. Vinyl monomers include methyl methacrylate, ethyl methacrylate (EMA), sec-butyl methacrylate (SBMA), styrene (St), acrylonitrile (An), ethylene-propylene-diene elastomer (EPDM), etc. The inorganic precursors include TEOS, tetramethyl orthosilicate (TMOS), titanium tetraisopropoxide (TIPO), aluminum tri-sec-butoxide (ASBO), etc.

As part of the same strategy, the sol–gel active functionalities can also be introduced to the polymer precursors via chemical transformation of the groups in preformed polymers. For example, poly[methyl methacrylate-co-3-(triethoxysilyl)propyl methacrylate] [P(MMA-ESMA)] precursors with designed number of trialkoxysilyl groups and molecular weights were prepared by the living group-transfer copolymerization of allyl methacrylate with methyl methacrylate using trimethylsilyl ketene acetal as initiator and tris(dimethylamino)-sulfonium bifluoride ($TASHF_2$) as catalyst, followed by hydro- silylation of the ester allyl groups in the polymer in the presence of the Speier's catalyst (H_2PtCl_6), as shown in Fig. 2 (Wei *et al.* 1990). Anhydride groups in vinyl polymers such as poly(styrene-co-maleic anhydride) can be converted to sol–gel active groups by treatment with 3-aminopropyltriethoxysilane (Zhou *et al.* 1998). Taking advantages of tremendous progress in the synthetic polymer chemistry, it is conceivable that polymer precursors with various architectures such as star and comb polymers, dendrimers, hyperbranched, and nanostructured polymers could be prepared to have sol–gel active functionalities along the main chains or as the end groups for ceramic modification.

Sol–gel active functional groups in polymer precursors also include those that can form extensive hydrogen bonds, ether (e.g., $\equiv$Si-O-C$\equiv$), ester (e.g., $\equiv$Si-O-CO-C$\equiv$) and other types of covalent bonds with ceramic sol–gel precursors. As examples, poly(methacrylic acid), poly(hydroxyethyl methacrylate), poly(vinyl alcohol), and poly(styrene sulfonic acid) can form hybrid materials with sol–gel ceramic matrices through ether, ester, and/or hydrogen bonding. The hydroxy groups in the sol–gel reaction intermediates and in the ceramic network (e.g., silanol group$\equiv$Si-O-H) may function as both the donor and acceptor of hydrogen bonds with a variety of functional groups such as hydroxyl, amine and carbonyl groups. Taking advantage of the acidity of the silanol proton, polymer–ceramic hybrid materials could also be obtained under proper conditions, based on ionic charge pairing between the ceramic component and ion-containing polymer chains.

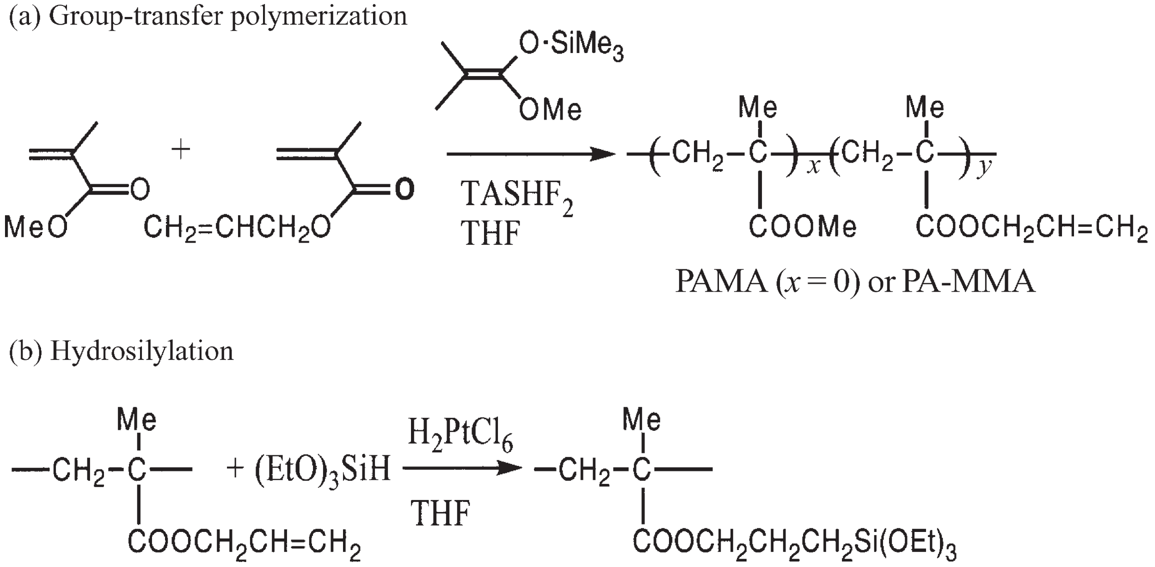


Figure 2
Synthesis of polyacrylate precursors via group-transfer polymerization.

In addition to vinyl polymers, polymer precursors can be prepared via step-growth or ring-opening polymerizations or copolymerizations. Among the examples are polyamides, polyimides, polyethers, polysiloxanes, polyesters, and epoxy polymers. In fact, polydimethylsiloxanes with sol–gel active -$SiMe_2$-OH end groups were used as polymer precursors in some pioneering works of synthesizing polymer–ceramic hybrid materials (e.g., Huang *et al.* 1985). The polysiloxane chains are covalently bonded to sol–gel silica matrix by forming the silyl ether linkage (i.e., $\equiv$Si-O-$SiMe_2$-). Like their counterpart in the vinyl polymer precursors, sol–gel active functional groups can be introduced to step-growth or ring-opening polymers by using the monomers that bear the sol–gel active groups in the polymerization or by post-polymerization transformations. Many polymers prepared by step-growth and ring-opening reactions already contain sol–gel active functionalities such as-OH, -COOH, -COCl, $\equiv$Si-OH, $\equiv$Si-Cl, etc., along the main chain or as the end groups. These polymers can be used directly in the sol–gel reactions with inorganic precursors to afford polymer-modified ceramics. It is important to emphasize that in the designing of any polymer precursors, the number of the sol–gel active functionalities per polymer chain should be great enough to overcome or compensate the thermodynamic repulsion between usually hydrophobic polymer and hydrophilic sol–gel ceramic components, in order to avoid macroscopic phase separation.

2.2 *Sol–Gel Reactions of Polymer Precursors*

In general, all the techniques for preparing pure sol–gel ceramics can be adopted or modified for the synthesis of polymer–ceramic hybrid materials. Typically, the sol–gel reactions of both the polymer and inorganic precursors are carried out in the presence of an acid (e.g., aqueous HCl) or a base catalyst (e.g., $NH_3 \cdot H_2O$) in a medium containing water and organic solvent, e.g., alcohols, tetrahydrofuran (THF) and *N*-methylpyrrolidinone (NMP), at room or elevated temperatures (usually $<100\,^{\circ}C$). By optimizing reaction conditions, hybrid materials in a form of bulk objects or thin films can be made transparent, monolithic, and free of cracks. The polymer–ceramic composition of the hybrid materials is governed by the amounts of polymer and inorganic precursors in the reactants. The chief purpose for using organic solvents is to dissolve both the polymer and inorganic precursors. The organic solvents are selected based on the following criteria: (i) they can dissolve the precursors; (ii) they have reasonable solubility in water; (iii) they have no or limited undesired reactions with the components in the system; and (iv) preferably, they are volatile compounds and can be removed readily by evaporation during the drying and aging of the gels. There is not much limitation in the choice of inorganic precursors, except for the consideration that the reactivity of the inorganic precursors should be comparable to that of the sol–gel active groups such as-$Si(OMe)_3$ in the polymer precursors. If the difference in reactivities is too high, there would be excessive self-condensation among either the polymer precursors or the inorganic precursors, leading to macroscopic phase separation. The high temperature densification or sintering of gels, as in the preparation of pure ceramics, is avoided to prevent thermal decomposition of the polymer component. It is known that acid-catalysis tends to yield more linear intermediates, while the base-catalysis tends to generate highly branched and, hence, globular intermediates in the sol–gel reactions before gelation. Therefore, the acid catalysts are often preferred, particularly when a product of excellent optical transparency is desired.

As an example, the synthesis of polymethacrylate–silica hybrid materials is achieved by the HCl-catalyzed sol–gel reactions of the trialkoxysilyl groups in the polymer precursor P(MMA–MSMA) with the inorganic precursor TEOS. The process consists of the following typical steps: (i) preparation of a solution (A) of the polymer precursor in an organic solvent (e.g., THF); (ii) HCl-catalyzed prehydrolysis of TEOS in a mixture of water and THF to give soln B; (iii) combining solutions A and B while stirring to afford a homogeneous solution (C); (iv) further reactions leading to gelation and formation of a homogeneous, transparent wet gel; and (v) aging and drying (by evaporating the solvents and byproducts, i.e., water and alcohol) to afford a homogeneous, transparent, and monolithic polymethacrylate–silica hybrid material. The purpose of the prehydrolysis is to form silanols and oligosilicates that can rapidly react with the polymer precursors before the occurrence of possible polymer–silicate phase separation in solution C. Various synthetic parameters, such as type and amount of catalysts and solvents, amounts of the polymers, inorganic precursors and water, reaction temperature, and rate of evaporation of the volatile, are optimized to assure optical transparency and integrity of the products.

A photochemically catalyzed sol–gel synthesis of polymer-modified ceramics was developed (Wei *et al.* 1994a). In this method, the catalyst for the sol–gel reactions is a photoacid, which is neutral in the dark and generates acid upon exposure to light. Among typical photoacids are diphenyliodonium chloride (DIC) and triphenylsulfonium hexafluoroantimonate (TSHF), both of which yield protonic acid under UV radiation. The acid thus generated catalyzes the sol–gel reactions. To prepare the polymethacrylate–silica hybrid materials, a mixture of the trialkoxysilyl-bearing polymer precursors, TEOS, THF, water, and a photoacid (e.g., DIC or TSHF) is exposed to UV radiation. After a certain time period, the initial organic/water phase separation disappears because of hydrolysis of alkoxysilyl groups and the system becomes a homogeneous solution. Subsequent gelation, aging, and drying steps are the same as in HCl-catalyzed process to yield homogeneous, transparent hybrid materials. As controls, no sol–gel reaction was observed for the system without the photoacid after exposure to the UV radiation or for the system with the photoacid that was kept in darkness for over 24 hours. This photochemical method is very general and can be readily extended to the synthesis of any type of polymer-modified sol–gel ceramics. It should find interesting applications in lithography, and in the coating and dental materials industries. By choosing appropriate photoacids, various light sources could be employed including sunlight.

The kinetics of the sol–gel reactions of mixed polymer and inorganic precursors have been studied to some extent. The presence of polymer precursors usually shortens the time required for the system gelation. In the case of poly[methyl methacrylate-co-3-(trimethoxysilyl)propyl methacrylate] [P(MMA-MSMA)] precursors, the higher the polymer content the shorter the gelation time. Such a trend was interpreted based on the fact that the polymer chain bears more reactive alkoxysilyl functional groups per molecule and, therefore, reaches gelation (i.e., formation of cross-linked network) sooner than the inorganic precursor TEOS. For the same reason, the polymer precursors containing more MSMA units have shorter gelation time than those containing fewer MSMA units (Wei *et al.* 1995a). Many multicomponent ceramic systems have also been investigated. For example, ternary P(MMA-MSMA)–SiO_2–Al_2O_3 and P(MMA–MSMA)–SiO_2–TiO_2 hybrid materials were synthesized via the sol–gel reactions of P(MMA-MSMA) with TEOS, TIPO, and/or ASBO. Alkoxides of different metal elements are known to have a wide range of reactivities toward hydrolysis. The titanium and aluminum precursors, TIPO and ASBO, react with water much faster than does TEOS, often resulting in inhomogeneous sol–gel products. Among the solutions to this problem are (i) to introduce of a chelating organic ligand into the reaction to reduce the hydrolysis rate of the highly reactive metal alkoxides; and/or (ii) to change the in addition sequence of the reactants so that the less reactive precursors will have longer time to react before the more reactive precursors are added.

2.3 Simultaneous Formation of Ceramics and Polymers

Despite the many great advantages of the sol–gel process, there are several severe, sometimes detrimental, drawbacks in this process, which have limited its technological applications. One of the inherent problems in conventional sol–gel process is exceedingly large volume shrinkage resulting from the removal of solvents and small molecular by-products of hydrolysis and polycondensation such as water and alcohol. The volume shrinkage can easily reach 75% or more. For example, the theoretical shrinkage from TEOS to silica without any solvent is 55% based on their bulk densities. Thus, the preparation of monolithic sol–gel materials with predictable and controllable shape and dimension becomes very difficult and time-consuming because of the high internal stress induced by the volume shrinkage.

Among a few strategies developed to overcome the problem is the simultaneous formation of ceramics and polymers for the preparation of polyacrylate–silica composites. One elegant design is to make the alkoxy groups in the inorganic precursors polymerizable such as $Si(OR)_4$, where R is acrylate monomer moiety, $CH_2{=}CMe–COO–$. The sol–gel reactions of $Si(OR)_4$ proceed simultaneously with a free-radical polymerization of the acrylate sol–gel by-product so that no small ROH molecules will be released from the system and the volume shrinkage is reduced significantly or nearly eliminated (Ellsworth and Novak 1991). One of the limitations of this method is that only a fixed, low ceramic content in the composites can be obtained. To increase the silica content, poly(silicic acid) was introduced to the synthetic system as the major silica source (Novak and Ellsworth 1993), which, however, often has an adverse effect on optical transparency of the products.

Another strategy is based on removing most of the volatile molecules from the reaction systems just before gelation (Wei *et al.* 1996). In this method, the sol–gel precursors are commercially available alkoxides such as TEOS and vinyl monomers such as 2-hydroxyethyl methacrylate (HEMA), which contain sol–gel active functional groups (e.g., OH in HEMA). First, acid-catalyzed sol–gel reactions of TEOS are allowed to proceed, in the presence of the vinyl monomer HEMA, until the system becomes homogeneous. The silanol (SiOH) groups in the silicate colloids or hydrolyzed TEOS then react with the OH groups in HEMA to form covalent linkage between the monomer and the silicate component. At this stage, a polymerization initiator, such as the thermal radical initiator benzoyl peroxide (BPO) or photoinitiator benzoin methyl ether (BME), is dissolved in the solution and the reaction system is placed under a dynamic vacuum to remove large amount of solvent (i.e., THF) and by-products of the sol–gel reactions. After exhaustive pumping, the resultant nonvolatile fluid can be cast in molds of desired size and geometry. Then the free-radical polymerization of the vinyl monomers can be initiated either by increasing the temperature when the thermal initiator is used or by exposing the system to light when the photoinitiator is used. The vinyl polymerization proceeds simultaneously with further sol–gel reactions to afford the organic–inorganic hybrid materials. Since the major portion of solvent and small molecular byproducts are removed before the gelation, the volume shrinkage from the casting stage to the final hybrid product is significantly reduced to about 6–20% at silica contents up to 47%.

Because of the low volume shrinkage, the shrinkage-induced stress is lessened and the probability of crack formation in the hybrid materials is, therefore, greatly reduced. The sample preparation time is greatly shortened, from a few weeks to about 12 hours or less. Furthermore, vinyl monomers other than HEMA can be introduced into the system before the initiation of the vinyl polymerization so that the hybrid materials containing a variety of organic polymers can be designed and prepared. Other methods, such as differential solvent extraction, for removing the volatile molecules before the system gelation have also been investigated.

3. General Characterization Techniques

In general, the structures of polymer-modified ceramics can be characterized by most of the techniques that are commonly used for characterization of polymers and ceramics. Solid-state nuclear magnetic resonance (NMR) spectroscopy (particularly ^{29}Si NMR for silica) has proven informative. Conventional solution NMR can be employed in monitoring the sol–gel reactions up to gelation. Electronic absorption spectroscopy (UV–vis) can be performed on a monolithic disk or film of the polymer–ceramic sol–gel materials to yield information about the structure and optical properties including transparency. Infrared spectroscopy (IR) is suitable for identifying characteristic bands of the polymer and ceramic components in the final products. From intensities of the IR absorptions, polymer–ceramic composition can be estimated. Thermogravimetric analysis (TGA) is particularly useful for determining compositions in the materials by decomposing all the organic polymer contents at high temperatures (e.g., $>700\,^{\circ}$C). For example, the silica contents in the vinyl polymer–silica hybrids measured from TGA generally agree well with the calculated values based on the stoichiometry of the starting materials. In addition, many other physical properties, such as bulk density, refractive index, and hardness, exhibit strong correlation with the polymer–ceramic compositions of the hybrid materials. Extraction of the materials with a good solvent for the polymer can be employed to access qualitatively if there is a strong, particularly a covalent, bonding between the polymer and ceramic components.

There have been many studies of morphology and microstructures of polymer-modified ceramics using techniques such as x-ray diffraction and transmission and scanning electron microscopy (TEM and SEM). Atomic force microscopy (AFM) has proven helpful in identifying polymer–ceramic phase separation, or lack of it, in the hybrid materials (Wei *et al.* 1998a). The internal porosity of the sol–gel materials is commonly measured by nitrogen adsorption–desorption isotherms (the BET method). Differential scanning calorimetry (DSC) serves as an important technique in characterizing the glass transition of the polymers in the ceramic matrices. Lack of a well-defined glass transition temperature (T_g) generally indicates that the polymer chains are uniformly distributed in the materials. Dynamic mechanical analysis (DMA) can be used not only to study the glass transition but also to evaluate mechanical properties of the materials. In principle, all the standardized tests (e.g., ASTM protocols) can be employed for evaluating the polymer-modified ceramics. The limitation for these tests is in the sample preparation. Because of the volume shrinkage and stress-induced cracking, as mentioned above, it is very difficult to prepare samples of large dimensions per specifications with good reproducibility. Therefore, a relatively small number of reports has been published on systematic evaluation of mechanical and other physical properties of the polymer–ceramic hybrid materials.

4. Properties and Potential Applications

Polymer-modified sol–gel ceramics are usually amorphous, transparent to visible light and, therefore, the scale of phase separation, if any, should be below the wavelengths of visible light (i.e., $<400\,$nm). In general, physiochemical properties of the materials are between those of pure polymers and pure ceramics as anticipated and can be controlled by varying the polymer–ceramic compositions. Polymer–ceramic hybrid materials can be made harder than polymers and less brittle than ceramics. Typically, as the ceramic content is increased, the glass transition temperature of the polymer measured with DSC and DMA increases, becomes less pronounced and eventually disappears, because of the increasing restriction of large-scale, cooperative segmental motions of the polymer chains. The thermal stability of some polymer components (e.g., polyacrylates) is enhanced as indicated by the higher thermal decomposition temperatures compared to those of pure polymers. Optical properties such as index of refraction can be continuously tuned, e.g., from ~ 1.45 for sol–gel silica to 1.49 for PMMA or 1.60 for polystyrene (Wei *et al.* 1992, 1993a). This has prompted investigations of these hybrid materials for applications in optical waveguides and circuitry (Xu *et al.* 1996). Nonlinear optical (NLO) properties have also been explored by introducing vinyl polymer precursors that contain NLO moieties with either the second ($\chi 2$) or third ($\chi 3$) harmonic generation into a sol–gel silica matrix. Taking advantage of the restricted motion of polymer chains in the ceramic matrices, thermal randomization of the $\chi 2$ moieties could conceivably be minimized if the formation of ceramic network occurs concurrently with poling under electric or magnetic field.

4.1 Mechanical and Surface Properties

Bulk density and hardness of polymer-modified ceramics increase with ceramic content. For example, Vickers hardness was found to increase drastically from $\sim 20\,\mathrm{kg\,mm^{-2}}$ for pure polystyrene to $\sim 120\,\mathrm{kg\,mm^{-2}}$ for polystyrene–silica hybrid materials at silica content of ~80% (Wei *et al.* 1993a). The value for pure sol–gel silica is $\sim 160\,\mathrm{kg\,mm^{-2}}$. The relationship between the hardness and the silica content becomes linear after the silica content reaches ~20% as dictated by percolation theory. Similarly, the hybrid materials exhibit superior scratch-resistance over polymers. There have been numerous studies of mechanical properties of polymer-modified ceramics by dynamic mechanical analysis (DMA) (e.g., Wei *et al.* 1998c). In general, the storage modulus (E') becomes higher and the loss factor ($\tan \delta$) is smaller as the ceramic content is increased. All the results suggest that polymers are stiffened and hardened by the introduction of ceramic components.

As a typical example of the standardized tests for mechanical properties, a comparative study using the standard compressive strength test (ASTM D695-91) was performed in the author's laboratory on P(MMA-MSMA)–silica hybrid materials prepared via the polymer precursor route and on conventional composites of PMMA with fused silica powder as filler (particle size: 8–17 μm). Figure 3 shows the stress–strain behavior of the hybrid materials (bottom) in comparison with that of conventional composites (top) at various silica contents (wt%) as indicated by the last number in the curve labels. Pure sol–gel (SG) silica and PMMA (Plexiglas$^{\mathrm{TM}}$) were also tested for comparison. It is clear that the behavior of the hybrids differs significantly from conventional composites. In the range of silica contents up to ~50%, the curves for the PMMA–silica composites are very similar. In sharp contrast, the Young's modulus, yield and compressive strength of the hybrid materials increase rapidly as the silica content is increased. At silica contents of 32% and higher, the yield stress increases drastically and the hybrid materials exhibit strain hardening, which is not observed for PMMA–silica composites with up to 50% silica filler. It was also noticed that during compressive tests most of the hybrid samples maintained their excellent optical transparency even after the yielding points, until the moment of implosion. However, the composite samples crazed (see *Polymer Glasses, Fracture of: Crazing*$^{\#}$) after yielding. These results could be interpreted based on the picture that in the hybrid materials there is no macroscopic phase separation. Similar linear correlations of modulus and silica content were observed with the standard flexural 3-point-bending test (ASTM C1161-90; C674-88) on poly(2-hydroxyethyl methacrylate) (PHEMA)–silica hybrid materials. The modulus increased from ~1.6 GPa for pure PHEMA to ~3.2 GPa for the hybrid samples with silica content of 34%.

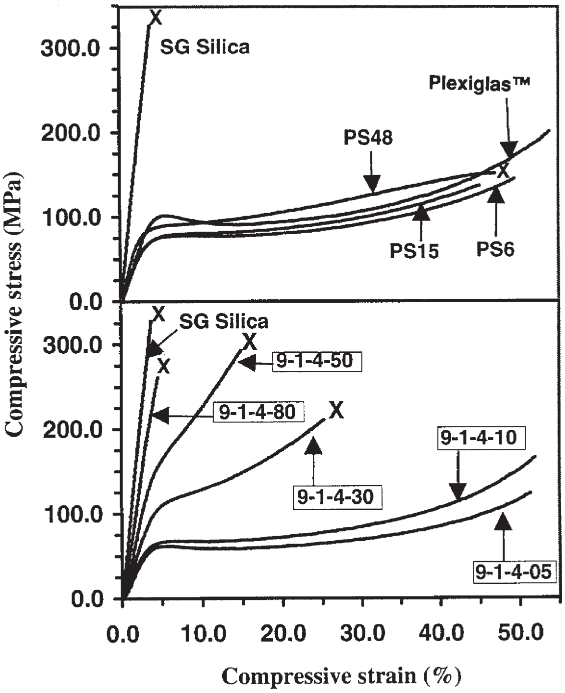


Figure 3
Stress–strain curves of conventional PMMA–silica composite (top) and P(MMA-MSMA)–silica hybrid materials with 10% MSMA in P(MMA-MSMA) precursor (bottom). The silica content (wt.%) of each sample is indicated by the last number on the curve label.

Based on their controllable mechanical properties, polymer-modified sol–gel ceramics can be designed for many applications. As examples, P(MMA-MSMA)–silica hybrid materials have been explored as a new generation of dental materials, which have chemical components similar to those of conventional dental composites (i.e., silica as fillers and polyacrylates as matrices) but no undesired macroscopic interfaces that could lead to mechanical failure and microleakage in the oral environment (Wei and Jin 1997). By tuning the hardness of the hybrid materials, hard coatings or pressure-sensitive coatings/capsules could also be obtained.

The polymer modification improves the impact strength of sol–gel ceramics significantly, particularly when the polymers of high flexibility and low T_g such as polydimethylsiloxanes are employed. Relative brittleness of vinyl polymer–ceramic hybrid materials was evaluated, however, qualitatively because of the difficulties in the preparation of samples for standardized tests. The brittleness of the PMMA–silica hybrid samples tends to decrease as the polymer content is increased. At the same polymer content, the materials prepared from the polymer precursors containing fewer sol–gel active trialkoxysilyl groups (i.e., $Y \ll X$ in Fig. 1) appear to be less brittle than those containing more trialkoxysilyl groups (e.g., $X = 0$). Such a trend can be rationalized by considering segmental motions of polymer chains that could absorb impact energy. The higher the trialkoxysilyl group content the greater the extent of cross-linking via the formation of Si–O bonds, which leads to lesser segmental motions through conformation changes of the polymer chains and, hence, more pronounced brittleness of the materials. For the same reason, the effect of molecular weight of the polymer components is usually less pronounced than that of extent of cross-linking on many mechanical properties. Though this article has emphasized the amorphous hybrid materials without macroscopic phase separation, it should be noted that for some applications the phase-separated materials are actually desired. For example, deformation or breakage of numerous small and weak interfaces could absorb externally applied energy or relax internal stress more uniformly to avoid catastrophic failure of the bulk material. Furthermore, polymer-modified ceramics with controllable degrees of crystallinity present interesting opportunities in the exploration of nanostructured materials.

Polymer-modified ceramics may exhibit designable surface properties because both polymers and ceramics, typically hydrophobic and hydrophilic, respectively, are exposed on the surface. For example, by controlling the compositions of PMMA–silica hybrid materials, the contact angles of water on the surface of the hybrid materials were found to be much greater than that of silica glass and to increase with the polymer content, approaching the value for pure PMMA. With consideration that the PMMA–silica hybrids should have better thermodynamic compatibility with polyacrylate resins, applications of the hybrid materials as filler to substitute for powdered silica in conventional dental composites have been investigated. Composites were prepared using both the PMMA–silica hybrid materials and the powdered silicas (with or without surface silane treatment) as fillers in a commercial 2,2-bis(*p*-2-hydroxy-3-methacryloxy-propoxyphenyl)propane (bis-GMA) resin system. The results of compressive tests showed that the samples with the hybrid fillers had elastic moduli and maximum strengths comparable to or higher than those with silane-treated silica fillers at the same silica content in the composites (Wei *et al* 1998b).

Introduction of sol–gel ceramic components to polymers can also significantly improve adhesion of polymers as coatings to inorganic substrates, which often contain surface hydroxy groups that can react with the sol–gel precursors. For example, polyaniline, an important electroactive polymer, could be employed as thin-film coatings in electrochromic, electrooptical, chemical/biological sensing, and electronic devices. However, the films exhibited poor adhesion to the substrates such as indium–tin oxide (ITO) and silica glasses, particularly in aqueous media, which severely limited the performance and lifetime of the devices. By modifying polyaniline solutions with relatively small amounts (e.g., 10 wt.%) of sol–gel precursors such as TEOS or trialkoxysilyl-containing polymethacrylates, adhesion of the solution-cast composite films to ITO and silica glasses was increased dramatically while the electrooptical properties of polyaniline remained satisfactory (Jang *et al.* 1996, Wei *et al.* 1995b).

4.2 Chemical Properties and Postsynthesis Modifications

Both polymer and ceramic components in polymer-modified sol–gel ceramics can undergo a variety of reactions to afford new materials of desired properties. As examples, the polyacrylonitrile component in hybrid silica materials was found to undergo pyrolysis in an inert atmosphere towards the formation of carbon fibers of micro- or nanostructures in the silica matrix (Wei *et al.* 1993b). Chemoselective hydrolysis of polyacrylonitrile component yields new water-soluble polyacrylic-modified silica materials via transformation of nitrile (-CN) groups to amide (-CONH_2) and carboxylate (-COO^-) groups (Wei *et al.* 1994b). Sulfonation of the polystyrene component in the polystyrene–silica hybrid materials resulted in new materials containing sulfonate groups with good cationic exchange properties (Wei *et al.* 1997). Similarly, chloromethylation can yield anionic exchange materials. Properly functionalized hybrid materials could also be used in chromatography, solid-phase protein synthesis, polyelectrolytes, catalysts, and membranes for bioseparations. The ceramic components can also undergo post-synthesis reactions. As examples, the Si-OH groups on the external or internal surfaces could be alkylated and silica content could be removed with reagents such as HF or strong bases.

The organic polymer content in polymer–ceramic hybrid materials could be removed by thermal or chemical decomposition to afford porous ceramic materials. For example, a PMMA–silica hybrid sample synthesized at room temperature was found to have a BET surface area of $\sim 0.1\,\text{m}^2\,\text{g}^{-1}$. After thermal treatment at ~400 °C in nitrogen, the polymer content was completely removed leaving a porous silica with a dramatically increased surface area of $\sim 630\,\text{m}^2\,\text{g}^{-1}$ and an average pore diameter of ~28 Å. By tailoring the nature and the composition of the polymers, it is possible to control the size and volume of the pores and/or microchannels in the porous inorganic materials (Wei *et al.* 1995a). These materials could also be made to contain other metal oxides such as titanium, vanadium, zirconium, etc., via multicomponent sol–gel process. Polymers such as poly(ethylene oxide) (PEO) have been employed to form PEO–ceramic sol–gel composites. Upon removal of PEO, meso- or nanoporous molecular sieves were obtained. There are many potential applications for these porous ceramic materials including catalysts/catalyst supports, reaction templates, host matrices for fabricating capacitors, and as size-selective membranes for construction of biosensor devices.

Polymers can also include many bioactive substances such as proteins, DNAs, RNAs, polysaccharides, etc. In this sense, there has been a large body of work on encapsulation or immobilization of biomolecules in sol–gel ceramics (e.g., Avnir *et al.* 1994, Wei *et al.* 1999). These biopolymer-modified ceramics have tremendous potential for applications in biocatalysis, biosensor, medical, and pharmaceutical industries.

5. Concluding Remarks

Polymer-modified ceramics have become an important area of research in both polymer and ceramic sciences since the 1980s. There has been a great deal of fundamental research on the synthesis, characterization, and properties of this new class of materials, including polymer–ceramic hybrid materials and polymer–ceramic molecular composites or nanocomposites. Many major synthetic strategies have been derived and a variety of new materials prepared. Advances in synthetic polymer chemistry allow for designing and preparing various monomers and polymers that are suitable for ceramic modification. Better understanding of sol–gel chemistry and ever-developing processing methodology make it possible to incorporate almost any conceivable polymer, natural or synthetic, into inorganic matrices at the molecular or nanometer scales. Basically, polymer-modified ceramics can be viewed as conventional polymer–ceramic composites at much shrunken scales (i.e., molecular composites or nanocomposites) or as single-phase (i.e., hybrid) materials. In terms of properties, polymer-modified ceramics have lived up to the expectations of their designers. These materials have physical, chemical, and biological properties continuously tailorable between those of pure polymers and pure ceramics in most cases. Such a versatile tunability would provide a unique and powerful means for materials design. Because of the mixing at the molecular or nanometer scales, polymer-modified ceramics such as polyacrylate–silica hybrid materials could be made transparent to visible light.

With controlled conditions, polymer-modified sol–gel ceramics could be prepared as films, coatings, fibers, powders, beads, and bulk objects of various shapes and dimensions. Numerous applications for these materials have been proposed and explored. So far the initial results are very promising. With further research efforts and investment, there will be enormous opportunities for the development of commercial products in both civilian and defense use, particularly coatings, optical, electronic, biotechnology, and dental materials industries.

Bibliography

Avnir D, Braun S, Lev O, Ottolenghi M 1994 Enzymes and other proteins entrapped in sol–gel materials. *Chem. Mater.* **6**, 1605–14

Brinker C J, Scherer G W 1990 *Sol–Gel Science, the Physics and Chemistry of Sol-Gel Processing*. Academic Press, San Diego, CA

Chujo Y, Ihara E, Kure S, Suzuki K, Saegusa T 1990 Block copolymer of 2-methyl-2-oxazoline with silica gel. An organic-inorganic hybrid polymer. *Polym. Prepr.* (*Am. Chem. Soc., Div. Polym. Chem.*) **31** (1), 59–60

Ellsworth M W, Novak B M 1991 Mutually interpenetrating inorganic–organic networks. New routes into nonshrinking sol-gel composite materials. *J. Am. Chem. Soc.* **113**, 2756–8

Huang H H, Orler B, Wilkes G L 1985 Ceramers: hybrid materials incorporating polymeric–oligomeric species with inorganic glasses by a sol–gel process. *Polym. Bull.* **14**, 557–64

Huang H H, Wilkes G L 1987 Structure–property behavior of new hybrid materials incorporating oligomeric poly(tetramethylene oxide) with inorganic silicates by a sol–gel process. *Polym. Bull.* **18**, 455–62

Jang J-W, Chen C, Gumbs R W, Wei Y, Yeh J-M 1996 Large-area electrochromic coatings: composites of polyaniline and polyacrylate–silica hybrid sol–gel materials. *J. Electrochem. Soc.* **143**, 2591–6

Jang C Y, Mark J E 1984 The effects of various catalysts in the in-situ precipitation of reinforcing silica in polydimethylsiloxane networks. *Makromol. Chem.* **185**, 2609–13

Klein L C, Francis L, DeGuire M R, Mark J E (eds.) 1999 *Organic/Inorganic Hybrid Materials II*. Materials Research Society Vol. 576. MRS, Warrendale, PA

Laine R M, Sanchez C, Giannelis E, Brinker C J (eds.) 1998 *Organic/Inorganic Hybrid Materials*. Materials Research Society Vol. 519. MRS, Warrendale, PA

Mackenzie J D, Ulrich D R (eds.) 1988 *Ultrastructure Processing of Advanced Ceramics*. Wiley, New York

Mark J E, Lee C Y-C, Bianconi P A (eds.) 1995 *Hybrid Organic–Inorganic Composites*. ACS Symposium Series 585. American Chemical Society, Washington, DC

Novak B M 1993 Hybrid nanocomposite materials—between inorganic glass and organic polymers. *Adv. Mater.* **2**, 337–9

Novak B M, Ellsworth M W 1993 "Inverse" organic–inorganic composite materials. High glass content nonshrinking sol–gel composites. *Mater. Sci. Eng. A* **162**, 257–64

Pope E J A, Asami M, Mackenzie J D 1989 Transparent silica gel–PMMA composites. *J. Mater. Res.* **4**, 1018–26

Schmidt H 1985 New type of non-crystalline solids between inorganic and organic materials. *J. Non-Cryst. Solids* **73**, 681–91

Wei Y, Bakthavatchalam R, Whitecar C K 1990 Synthesis of new organic–inorganic hybrid glasses. *Chem. Mater.* **2**, 337–9

Wei Y, Jin D 1997 A new class of organic–inorganic hybrid dental materials. *Polym. Prepr. (Am. Chem. Soc., Div. Polym. Chem.)* **38** (2), 122–3

Wei Y, Jin D, Brennan D J, Nieves Rivera D, Zhuang Q, DiNardo J, Qiu K 1998a Atomic force microscopy study of organic–inorganic hybrid materials. *Chem. Mater.* **10**, 769–72

Wei Y, Jin D, Wei G, Yang D, Xu J 1998b Novel organic–inorganic chemical hybrid fillers for dental composite materials. *J. Appl. Polym. Sci.* **70**, 1689–99

Wei Y, Jin D, Yang C, Kels M C, Qiu K Y 1998c Organic–inorganic hybrid materials: relations of thermal and mechanical properties with structures. *Mater. Sci. Eng. C* **6**, 91–8

Wei Y, Jin D, Yang C, Wei G 1996 A fast and convenient method for the preparation of monolithic hybrid sol-gel materials without significant volume shrinkage. *J. Sol-Gel Sci. Technol.* **7**, 191–201

Wei Y, Wang W, Jin D, Yang D, Tartakovskaya L 1997 Synthesis of sulfonated polystyrene–silica hybrids and their applications as ion-exchange materials. *J. Appl. Polym. Sci.* **64**, 1893–902

Wei Y, Wang W, Yeh J-M, Wang B, Yang D, Murray J K Jr. 1994a Photochemical synthesis of polyacrylate–silica hybrid sol–gel materials catalyzed by photoacids. *Adv. Mater.* **6**, 372–4

Wei Y, Wang W, Yeh J-M, Wang B, Yang D, Murray J K Jr., Jin D, Wei G 1995a Vinyl polymer-modified hybrid materials and photoacid-catalyzed sol–gel reactions. In: Mark J E, et al. (ed.) *Hybrid Organic–Inorganic Composites.* ACS Symposium Series 585. American Chemical Society, Washington, DC, pp. 125–41

Wei Y Xu J, Jin D, Lin M, Feng Q 1999 "Room with a view": novel immobilization of enzymes in mesoporous host materials. *Proc. North American Catalysis Society Conf.* (*NAM'99 Boston*), North American Catalysis Society, Boston, MA, p. 14

Wei Y, Yang D, Bakthavatchalam R 1992 Thermal stability and hardness of new polyacrylate–SiO_2 hybrid sol–gel materials. *Mater. Lett.* **13**, 261–6

Wei Y, Yang D, Tang L, Hutchins M K 1993a Synthesis, characterization and properties of new polystyrene–SiO_2 hybrid sol–gel materials. *J. Mater. Res.* **8**, 1143–52

Wei Y, Yang D, Tang L 1993b Synthesis of new polyacrylonitrile–silica hybrid sol–gel materials and their thermal properties. *Makromol. Chem., Rapid Commun.* **14**, 273–8

Wei Y, Yang D, Wang W, Tang L 1994b Synthesis of water-soluble-polyacrylics modified silica sol-gel composite materials. *Chem. Mater.* **6**, 1737–41

Wei Y, Yeh J-M, Jin D, Jia X, Wang J, Jang G-W, Chen C, Gumbs R W 1995b Composites of electronically conductive polyaniline with polyacrylate–silica hybrid sol–gel materials. *Chem. Mater.* **7**, 969–74

Wilkes G L, Huang H, Glaser R H 1990 New inorganic–organic hybrid materials through the sol–gel approach. In: Ziegler J M, Fearon F W G (eds.) *Silicon-Based Polymer Science,* Advances in Chemistry Series 224. American Chemical Society, Washington, DC, pp. 207–26

Xu C, Eldada L, Wu C, Norwood R A, Shacklette L W, Yardley J T, Wei Y 1996 Photoimageable, low shrinkage organic–inorganic hybrid materials for practical multimode channel waveguides. *Chem. Mater.* **8**, 2701–3

Zhou W, Dong J-H, Qiu K Y, Wei Y 1998 Preparation and properties of poly(styrene-co-maleic anhydride)/silica hybrid materials by the in situ sol–gel process. *J. Polym. Sci., Part A: Polym. Chem.* **36**, 1607–13

Y. Wei
Drexel University, Philadelphia, Pennsylvania, USA

Polymer–Nonsilica Ceramic Nanocomposites

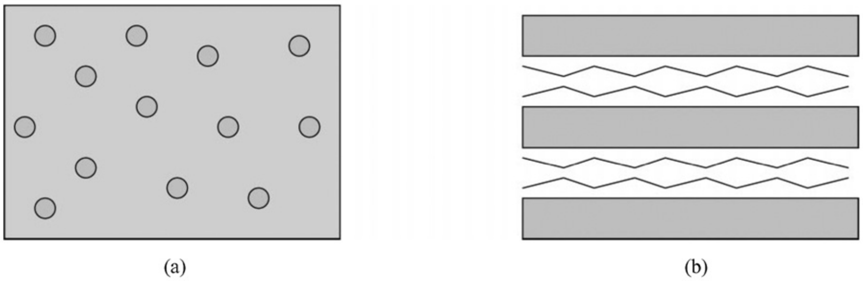
A nanocomposite comprises at least one phase with dimensions in the nanometer regime (1–100 nm) for its combination of two or more phases (Gonsalves and Chen 1996a, 1996b, Lerner and Oriakhi 1997). The form of the nanophase can be grains, clusters, or fibers with a nanoscale diameter, or layers with a thickness in the nanometer range (Dagani 1992). The high surface-to-volume ratio of the nanophase in nanocomposites has made them extraordinary materials with unusual properties including electrical, optical, magnetic, gas barrier, thermal, and mechanical properties. They have shown promise in important applications such as digital information storage, color imaging and printing, medical diagnostics, leak-free sealing, detection sensing devices, magnetic refrigeration (Dagani 1992), ferrofluids (Dagani 1992, Meldrum *et al.* 1992), lithium rechargeable batteries (Koene and Nazar 1996), light-emitting diodes (Alivisatos 1996), gas separation membranes (Hu *et al.* 1997), and structural materials (Wang and Mark 1987, Mark and Wang 1988). Although existing nanocomposites contain only two phases, those having more phases will emerge to satisfy other needs.

One class of nanocomposites with fascinating properties is polymer–ceramic nanocomposites. They often possess desirable combinations of properties from both polymer and ceramic components. Although the term 'polymer' has an exact definition, the term 'ceramic' has several approximate definitions. Based on these definitions (Brinker and Scherer 1990, Riedel and Dressler 1996, Calvert 1996, Zhang and Gonsalves 1994), metal or metalloid oxides, nitrides, carbides, borides, sulfides, silicides, phosphides, and carbonates may reasonably be considered as ceramics. However, ceramics are inorganic materials, both crystalline and noncrystalline in nature.

Properties of polymer–ceramic nanocomposites mainly depend on the size, size distribution, shape, and homogeneous dispersion of the nanophase, and morphology, relative amounts, orientation, and interactions of both polymer and ceramic phases. The design and preparation of a polymer–ceramic nanocomposite with desired properties are a challenge. Therefore, it is important to explore the synthesis routes of such nanocomposites and their resulting properties. This article addresses the synthesis and the properties of polymer–nonsilica ceramic nanocomposites.

1. Host–Guest Behavior

There are two main types of polymer–ceramic nanocomposite, based on their host–guest architecture. The first type of nanocomposite consists of nanoscale



(a) (b)

Figure 1
Structures of polymer–ceramic nanocomposites in which (a) ceramic nanoparticles reside in a polymer host matrix and (b) polymeric guests are intercalated in nanoscale galleries between layered ceramic hosts.

ceramic guests in a polymer host matrix (Fig. 1(a)). In such an intimate environment, nanoparticles can benefit from the polymer host matrix or the polymer matrix can benefit from the nanoparticle guests, or both, to improve their properties. Nanoparticles themselves reveal exceptional mechanical, electrical, and optical properties. For instance, ceramic semiconductor nanoparticles such as CdS (Alivisatos 1996) exhibit remarkable optoelectronic properties that cannot be obtained either from individual molecules or bulk semiconductors. Such optoelectronic properties arise from the quantum confinement of electrons in the nanoparticles. In other words, electrons in semiconductor nanoparticles, or so-called quantum dots, occupy a small number of discrete energy levels as is the case with the "particle in a box" concept. These electrons can absorb energy from electricity or sunlight and emit energy as electricity or light. In order to take advantage of such nanoparticles, it is very important to overcome the main obstacles in obtaining unagglomerated nanoparticles with desired uniform sizes. One way to achieve this is to fabricate polymer–ceramic nanocomposites in which there is a homogeneous dispersion of unagglomerated uniform particles throughout a processable polymer matrix. The benefits from ceramic nanoparticle guests for their polymeric hosts in a polymer–ceramic nanocomposite include improved mechanical and gas permeability properties. By introducing an inorganic phase into a polymer, the mobility of polymer chain segments can be hindered and chain packing can be inhibited, resulting in an increase in gas separation capabilities (Hu *et al.* 1997). Also, ceramic nanoparticles in a polymer matrix can act as reinforcing fillers improving mechanical properties (Wang and Mark 1987, Mark and Wang 1988) (see *Elastomers: Ceramic-modified**). Furthermore, the electrochemical properties of polymer electrolytes can be improved by incorporating ceramic nanoparticles into the polymer phase (Krawiec *et al.* 1995).

The second type of polymer–ceramic nanocomposite consists of a ceramic host matrix having nanosize cavities and polymeric guests residing in these cavities (Fig. 1(b)). These polymeric guests benefit from the ceramic host. Conjugated polymers of this kind possess high electrical conductivity and large nonlinear susceptibility. Such important properties of conjugated polymers can be improved by enhancing the degree of chain orientation. Intercalation of a conjugated polymer in nanoscale spaces in ordered ceramic crystalline structures gives rise to an oriented polymer and thereby results in a nanocomposite having a conjugated polymer with improved electrical conductivity and nonlinear optical properties (Nazar *et al.* 1992).

2. Synthesis of Nanocomposites of Nonsilica Ceramic Guests and Polymeric Hosts

2.1 Blending

Agglomeration of nanoparticles during the preparation of polymer–ceramic nanocomposites is one of the main disadvantages of the conventional blending method. However, advances in synthesis of nanocomposites have not abandoned the blending process. Polyethylene oxide (PEO)/Al_2O_3 nanocomposites (Krawiec *et al.* 1995) have been prepared by using a homogeneous mixture of PEO, Al_2O_3 nanoparticles and Li_2F_4 in acetonitrile. The agglomeration of Al_2O_3 particles is prevented by agitating the mixture. Composite films prepared by casting the resulting mixture and evaporating the acetonitrile have shown homogeneously dispersed alumina nanoparticles in the polymer matrix. The ionic conductivity of the resulting composites is increased by an order of magnitude compared to the same composites having alumina microparticles, also at ambient temperature.

Polymethylmethacrylate/CdS nanocomposites having significantly high third-order optical nonlinearity have been prepared by uniformly dispersing phenyl–CdS nanoparticles with narrow size distributions in the PMMA matrix (Yanagida *et al.* 1990). Films of these composites have been formed by mixing a solution of PMMA in pyridine and a solution of phenyl–CdS in pyridine and casting the resulting mixture. Surface modification of CdS with phenyl capping prevents particle agglomeration.

The use of polyimides in microelectronic applications is limited because of their low thermal conductivity and high thermal expansion coefficient. Incorporating aluminum nitride (AlN) nanoparticles into the polymer matrix can diminish such drawbacks with polyimide. In this process, the agglomeration of AlN particles is successfully prevented. Agglomerated AlN nanoparticles were first partially deagglomerated by suspending in *N*-methylpyrrolidone. A solution of polyimide precursor, polyamic acid (PAA), in *N*-methylpyrrolidone was then added to this suspension and the resulting mixture agitated for further deagglomeration. It is thought that the deagglomeration of AlN nanoparticles is due to several factors, such as electrostatic stabilization of particles by the polar organic solvent *N*-methylpyr5rolidone and the polyelectrolyte polymer PAA, interpenetration of polyamic acid chains into the loosened structure of agglomerates, and mechanical agitation. Polyimide/AlN nanocomposite samples have been prepared from the resulting homogeneous suspension by casting (Gonsalves and Chen 1994) followed by thermal imidization or rapid precipitation with triethylamine followed by thermal imidization and then hot pressing (Gonsalves and Chen 1995, Gonsalves *et al.* 1997). These composites were a homogeneous dispersion of nanosized AlN particles in the polymer matrix. Thermal conductivity and mechanical properties such as hardness and Young's modulus were increased and the thermal expansion coefficient was decreased for the hot-pressed composite samples with increasing amount of AlN present.

2.2 *In Situ Precipitation*

In situ precipitation is a very efficient method of distributing ceramic nanoparticles homogeneously throughout the polymer matrix without particle agglomeration. Another advantage of this method is its ability to control particle sizes. *In situ* generation of ceramic metal oxide particles in a polymer matrix is often achieved by the sol–gel process. Sol–gel technology (Brinker and Scherer 1990) offers an attractive alternate route to prepare ceramics at, or near, ambient temperatures unlike the traditional high-temperature sintering method. Using this process, the composition and the microstructure of the ceramic can be controlled at the molecular level and the desired shape of the material can be achieved. In addition, high purity of the sol–gel precursors yields higher ceramic purity relative to the ceramics traditionally made by melting minerals. The sol–gel process involves the hydrolysis of metal alkoxide precursors such as organosilicates, -titanates, and -aluminates, followed by condensation of the multihydroxy compounds produced (see *Sol–Gel Reactions: Theory*[#]). These hydrolysis and condensation reactions can be catalyzed either by acids or bases (see *Polymer–Ceramic Nanocomposites: Catalysts**). Easy and proper mixing of polymer into liquid sol–gel precursors and the low processing temperatures which preserve the organic polymer from degradation are among the benefits of the sol–gel process in fabricating polymer–ceramic nanocomposites.

Although silica is used as an efficient reinforcing filler for polydimethylsiloxane (PDMS), it may cause an adverse effect on high-temperature stability of PDMS. In fact, the acidic SiOH groups on the silica particle surface can cleave the PDMS chain. Such an effect on PDMS has not been observed when titania is used as the reinforcing filler and the mechanical properties of the resulting composite have been improved. These elastomeric PDMS/TiO_2 nanocomposites (Wang and Mark 1987) have been prepared via the sol–gel process by first swelling the end-linked PDMS networks with the sol–gel precursor *n*-propyltitanate. Swelled networks are then immersed in an aqueous solution of the catalyst (HCl, NH_4OH, or $(C_2H_5)_2NH$) to precipitate the TiO_2 *in situ*. NH_4OH has proven to be the most efficient catalyst. Moduli of the resulting dried composites have shown an increase with increasing amount of TiO_2.

Similarly, alumina as a reinforcing filler has been *in situ* precipitated in end-linked PDMS networks employing aluminum tri-*sec*-butoxide as the sol–gel precursor and HCl or NH_4OH as the catalyst (Mark and Wang 1988). Modulus and the ultimate strength of the resulting elastomeric PDMS/Al_2O_3 composites were increased, while their ultimate elongation and toughness decreased with increasing amount of Al_2O_3.

Aromatic poly(amide-imide)s possess a combination of properties such as high thermal stability, chemical resistance, and gas permeability and permselectivity from polyimides and high mechanical properties from polyamides. In an attempt to improve the gas permeability properties of poly(amide-imide)s further, poly(amide-imide)/TiO_2 nanocomposites have been prepared by *in situ* precipitation of TiO_2 nanoparticles using sol–gel technology (Hu *et al.* 1997). First, a sol prepared by partially subjecting the precursor tetraethyltitanate to sol–gel reactions was mixed with a solution of the poly(amide-imide). Then the resulting homogeneous mixture was cast and the films formed upon drying were heat-treated for further condensation of the inorganic precursor. These composites have shown improved high permselectivities for selected gas pairs even at very low titania levels in the membrane. TiO_2 particles seem to have specific interactions with gases such as CO_2 and H_2.

In one study, a homogeneous dispersion of size-controlled TiO_2 nanoclusters (1–1.5 nm) in a polyimide matrix were obtained by sol–gel *in situ* precipitation and chemically attaching the inorganic clusters to the specific sites on the polymer backbone (Nandi *et al.* 1991). Interestingly, this sol–gel precipitation has been supported by the chemical conversion of the polyimide precursor polyamic acid. During the preparation, first the polyamic acid is formed from a solution of dianhydride and diamine in the presence of the sol–gel precursor tetraethyltitanate. Carboxylic groups on the polyamic acid chains act as binding sites for the titania precursor. The films cast from the resulting solution are heat-treated, at which the imidization occurs. Water released from the imidization initiates the sol–gel processing of pre-bound titania precursor, giving rise to titania nanoclusters. The high glass-transition temperature of polyimides may also have been a factor in preventing the cluster agglomeration, since it decreases the cluster mobility inside the polymer matrix.

In situ precipitation of ceramic metal sulfide particles in a polymer host matrix is generally achieved by exposing a mixture of polymer and the ceramic precursor to H_2S gas. For instance, polystyrene phosphonate diethylester (PSP)/cellulose acetate (CA)/CdS polymer–ceramic nanocomposites have been prepared by *in situ* precipitation of CdS semiconductor nanoparticles in PSP/CA membranes (Yuan *et al.* 1992). Initially, $Cd(NO_3)_2$ was incorporated into PSP/CA membranes either by floating PSP/CA membranes in a $Cd(NO_3)_2$ solution or casting followed by drying of a mixture of PSP, CA, and $Cd(NO_3)_2$ solutions. CdS particles have been generated subsequently by exposing the resulting films to H_2S gas.

In situ precipitation of ceramic nanoparticles accompanied by *in situ* copolymerization has been successfuly employed to obtain a narrow particle size distribution and a homogeneous dispersion of particles in methacrylate–styrene/PbS nanocomposites (Gao *et al.* 1994). Because of their contents, these composites have shown a significantly higher third-order nonlinear susceptibility. The first step in synthesizing methacrylate–styrene/PbS composites involves the formation of a lightly crosslinked microgel containing Pb^{2+} using a solution of lead methacrylate ($Pb(MA)_2$) in the presence of the initiator azobisisobutyronitrile (AIBN). Subsequently, this microgel is dissolved in styrene and then treated with H_2S gas. Finally, the resulting styrene organosol is subjected to further polymerization using AIBN to obtain a transparent solid. In the copolymerization with styrene, $Pb(MA)_2$ acts as a crosslinking agent.

3. Synthesis of Nanocomposites of Polymeric Guests and Nonsilica Ceramic Hosts

3.1 In Situ *Intercalative Polymerization*

A ceramic structure in which polymer molecules reside is quite remarkable. It is difficult to insert polymer chains into the micro- or nanoscale cavities in ceramic structures by diffusion. Therefore, the direct insertion of polymer into the ceramic structure is not an effective method to fabricate nanocomposites with polymeric guests and ceramic hosts. One successful method to fabricate such composites is *in situ* intercalative polymerization. One strategy involved in this process is the insertion of small monomer molecules instead of polymer chains into those galleries or pores in the ceramic structure followed by *in situ* polymerization of inserted monomers. Some details are given in *Clay-based Polymer Nanocomposites**.

Ceramic hosts with oxidizing ability can initiate the polymerization of monomers such as aniline and pyrole and consequently the polymerization occurs simultaneously with the intercalation. In non-oxidizing hosts, *in situ* polymerization of these monomers is assisted by external oxidizing agents, such as oxygen, after the intercalation. In some cases, oxidizing cations have been incorporated into the ceramic structure prior to the monomer intercalation.

Polyaniline/$HMWO_6$ ($M = Nb$, Ta) nanocomposites have been prepared by *in situ* intercalative polymerization (Koene and Nazar 1996). Powdered $HMWO_6.nH_2O$ is stirred with aniline for several days at 100 °C in a sealed tube *in vacuo*. At this stage, the strong Brønsted acidity of these layered oxide hosts promotes the intercalation of the weakly basic aniline. The inserted monomer is then polymerized *in situ* at 130 °C in air for two days. The resulting composites have shown an increase in lithium ion mobility within their structure compared to $HMWO_6$. This is attributed to the fact that the presence of the conjugated polymer polyaniline reduces the interaction of lithium ion with the $HMWO_6$ structure. These results suggest that polyaniline/$HMWO_6$ nanocomposites are potential candidates for cathode materials in lithium rechargeable batteries in which the lithium ion mobility is an important issue.

Since $V_2O_5.nH_2O$ is an oxidizing host for aniline, it has to be reduced before the controlled polymerization is initiated in the preparation of polyaniline/$V_2O_5.nH_2O$ nanocomposites (Liu *et al.* 1993). In this synthesis, V_2O_5 xerogel is mixed with a solution of anilium iodide in CH_2Cl_2. Oxidizing ability of the V_2O_5 ordered lamellar structure is diminished by the iodide, a better reducing agent than anilium. The intercalated aniline is then subjected to oxidative polymerization in air. The resulting composites have shown a slightly enhanced electrical conductivity.

3.2 Simultaneous Insertion and Crystal Growth

In another interesting approach, polymer–ceramic nanocomposite structures in which the polymers are molecularly run through the ceramic crystalline phases have been fabricated (Messersmith and Stupp 1992). In this process, inorganic ceramic crystals have been nucleated and grown in homogeneous solutions containing polymers as co-solutes. Consequently, polymer chains have been entrapped in sites such as grain boundaries or other two-dimensional defects, nanopores, lattice channels, or interlamellar spaces of the ceramic structure. Synthesis of these architectures in which the ceramic phase is tetracalcium aluminate and the organic polymer is poly(vinyl alcohol), poly(dimethyldiallyl ammonium chloride), or poly(dibutyl ammonium iodide) has been reported.

3.3 Exfoliation–Adsorption

Layered ceramic structures can be disassembled in a suitable solvent and the resulting dispersed single layers can be restacked by applying the necessary chemical and physical conditions. This experience has been successfully exploited to intercalate polymer guests in layered ceramic hosts giving rise to polymer–ceramic layered nanocomposites. In this method, layers of the ceramic hosts are first exfoliated into a colloidal form. A solution of the polymer is then added to this colloid so that the polymer adsorbs on the dispersed single layers. The resulting layers are subjected to spontaneous or induced flocculation. Polyethylene oxide/MoS_2 layered nanocomposites have been prepared using this strategy (Lemmon and Lerner 1994), and X-ray diffraction results have suggested the presence of polymer double layers within the MoS_2 galleries.

3.4 Exfoliation–Adsorption Followed by In Situ *Conversion*

Some polymers are unable to dissolve in solvents that are compatible with the colloidal form of the desired layered ceramics. In addition, some may not be formed by oxidative polymerization. Accordingly, the polymer–ceramic nanocomposites associated with such polymers cannot be synthesized through exfoliation–adsorption or *in situ* intercalation polymerization methods. However, this has been accomplished by the use of soluble derivatives of these polymers as precursors. This synthesis route involves the intercalation of soluble precursor polymer into the layered ceramic host matrix by the exfoliation–adsorption method and subsequent thermal or chemical conversion of this precursor to form the desired polymer *in situ*. Examples of polymer–ceramic nanocomposites prepared through this process are poly(*p*-phenylenevinylene) (PPV)/MoO_3 composites (Nazar *et al.* 1992). In the preparation of PPV/MoO_3 composites, an ionomer, poly(*p*-xylene-α-dimethylsulfonium chloride), is used as the precursor to PPV and a solution of this ionomer is added to a colloidal dispersion of $[A(H_2O)_2]_{0.25}MoO_3$ ($A = Na$, Li) in water. At this stage, the high molecular weight ionomer has been inserted into the interlamellar gap in the ceramic structure as linear chain molecules. Finally, the ionomer is converted to PPV by a thermal treatment. The opposite nature of the charges of the precursor polymer and the reduced molybdenum oxide plates can also facilitate the orientation of polymer chains in the host structure. The resulting composites show well-ordered stacks of polymer/ceramic layers. Although both PPV and MoO_3 are electrical insulators, PPV/MoO_3 thin films have shown increased electrical conductivity at room temperature. Thermally induced charge transport behavior of these materials suggests their suitability for semiconductor materials.

4. Biomimetic Approach

Nature has an incredible ability to design and biologically construct polymer–ceramic nanocomposites that are used to fulfill its needs. Such biocomposites include bones, teeth, and shells. They are formed by biomineralization in which the mineral guests are *in situ* precipitated within preorganized polymeric hosts. These hosts can be extracellular (intercellular) polymeric matrices, or self-assembled protein cages or lipid vesicles (Mann 1993). The well-ordered assembly of organic and inorganic constituents and their controlled sizes, shapes, spatial distributions, and chemical and physical properties are responsible for the remarkable properties of biocomposites.

Bone is a collagen polymer matrix reinforced by nanosized ribbon-shaped crystals of hydroxyapatite ceramics (Burdon *et al.* 1995). The mechanical properties of these biological composites are superior to those of synthetic nanocomposites. In fact, they can maintain a significant toughness with high modulus and tensile strength, unlike synthetic nanocomposites (Calvert 1996). Ferritin, which functions as an iron store in the liver, is a spherical-shaped shell (apoferritin) filled with a hydrated iron oxide ferrihydrate ($5Fe_2O_3.9H_2O$) core (Meldrum *et al.* 1992). External and internal diameters of the ferritin micelle are approximately 12 nm and 8 nm, respectively. Host–guest architecture of both bone and ferritin are formed by biomineralization (Mann 1993).

The mimicry of Nature's architectures and the mechanisms through which they are made can result in better synthetic polymer–ceramic nanocomposites. Several polymer–nonsilica ceramic nanocomposites have been prepared by following the guidelines of the formation of biocomposites with polymeric host matrices. The already mentioned *in situ* precipitated polymer–nonsilica ceramic nanocomposites can fall into this category. A nanocomposite containing polypropylene matrix and elongated titania particles has been prepared by extruding the melt blended mixture of polypropylene and titanium *n*-butoxide followed by sol–gel processing of the alkoxide (Burdon *et al.* 1995). Proper interfacial bonding between the titania particles and polymer phase was achieved by using a titanate coupling agent (see *Polymer–Ceramic Nanocomposites: Interfacial Bonding Agents**). Although the resulting composites were stiff, their strength was poor. Calcium carbonate–chitosan nanocomposites prepared using the biomimetic approach have also been reported (Zhang and Gonsalves 1994).

There have been several successful attempts in which polymer–nonsilica ceramic nanocomposites were synthesized, mimicking the biomineralization in protein cages (Meldrum *et al.* 1992, Chan *et al.* 1995). The goal was to generate nanoscale magnetic particles. The usefulness of these protein cages is that they can act as reaction cages to control the nucleation and growth of these ceramic nanoparticles. In one study, ferrimagnetic iron oxide magnetite Fe_3O_4 nanoparticles were formed using horse spleen ferritin cages (Meldrum *et al.* 1992). The first step of this process involved the removal of existing iron core from ferritin to form apoferritin by dialysis. Reconstruction of the apoferritin was achieved by slow addition of iron(II) solution to the apoferritin solution under slow oxidative conditions. The resulting magnetic protein, ferrofluid or magnetoferritin, contained uniform sized (6 nm) spherical Fe_3O_4 single crystals as the core. In a similar approach, nanoscale iron(III) hydroxide clusters have been synthesized employing horse spleen ferritin cages (Chan *et al.* 1995).

5. *Concluding Remarks*

Polymer–nonsilica ceramic nanocomposites comprise a combination of optimal properties from both polymer and ceramic phases, or even exclusive properties that are not related to either of these phases (from quantum confinement effects). These properties are favored by the intimate host–guest relationship of the constituents and the elegant architectures of the nanocomposites. A number of new structural designs and synthesis strategies have been successfully employed to obtain novel polymer–nonsilica ceramic nanocomposites with desirable properties. The mimickry of the structures of natural polymer–ceramic nanocomposites and their synthesis routes cannot be underestimated in this regard. Important applications such as communications, computing, and biomedical devices will potentially be developed in the light of such polymer–nonsilica ceramic nanocomposites.

Bibliography

Alivisatos A P 1996 Semiconductor clusters, nanocrystals, and quantum dots. *Science* **271**, 933–7

Brinker C J, Scherer G W 1990 *Sol–gel Science: The Physics and Chemistry of Sol–Gel Processing*. Academic Press, San Diego, CA

Burdon J, Szmania J, Calvert P 1995 Organic–inorganic hybrids with a crystalline polymer matrix. In: Mark J E, Lee C Y-C, Bianconi P A (eds.) *Hybrid Organic–Inorganic Composites*. American Chemical Society, Washington, DC, pp. 6–18

Calvert P 1996 In: Mann S (ed.) *Biomimetic Materials Chemistry*. VCH, New York

Chan P, Chua-anusorn W, Nesterova M, Sipos P, St. Pierre T G, Ward J, Webb J 1995 Organic–inorganic nanocomposites based on iron-containing clusters and biomolecules. *Aust. J. Chem.* **48**, 783–92

Dagani R 1992 Nanostructured materials promise to advance range of technologies. *Chem. Eng. News.* **72** (47), 18–24

Gao M, Yang Y, Yang B, Bian F, Shen, J 1994 Synthesis of PbS nanoparticles in polymer matrices. *J. Chem. Soc., Chem. Commun.* 2779–80

Gonsalves K E, Chen X 1994 Polyimide–AlN nanocomposites: synthesis and characterization. *Polym. Mater. Sci. Eng.* **71**, 512–3; Abstracts of papers of ACS **208**, 276

Gonsalves K E, Chen X 1995 Synthesis and investigation of the thermal behavior of high-loaded AlN/polyimide nanocomposites. *Polym. Mater. Sci. Eng.* **73**, 285–6; Abstracts of papers of ACS **210**, 151

Gonsalves K E, Chen X 1996a In: Salamone J C (ed.) *Polymeric Materials Encyclopedia*. CRC Press, Boca-Raton, FL

Gonsalves K E, Chen X 1996b Studies on the structure and properties of ceramic/polymer nanocomposites. In: Coltrain B K, Sanchez C, Schaefer D W, Wilkes G W (eds.) *Better Ceramics Through Chemistry VII: Organic/Inorganic Hybrid Materials*. Materials Research Society, Pittsburgh, PA, pp. 55–65

Gonsalves K E, Chen X, Baraton M-I 1997 Mechanistic investigation of the preparation of polymer/ceramic nanocomposites. *Nanostructured Mater.* **9**, 181–4

Hu Q, Marand E, Dhingra S, Fritsch D, Wen J, Wilkes G 1997 Poly(amide-imide) TiO_2 nano-composite gas separation membranes: fabrication and characterization. *J. Membr. Sci.* **135**, 65–79

Koene B E, Nazar L F 1996 Synthesis and electrochemical lithium insertion in polyaniline/$HMWO_6$ ($M = Ta$, Nb) nanocomposites. *Solid State Ionics* **89**, 147–57

Krawiec W, Scanlon Jr. L G, Fellner J P, Vaia R A, Vasudevan S 1995 Polymer nanocomposites: a new strategy for synthesizing solid electrolytes for rechargeable lithium batteries. *J. Power Sources* **54**, 310–5

Lemmon J P, Lerner M M 1994 Preparation and characterization of nanocomposites of polyethers and molybdenum disulfide. *Chem. Mater.* **6**, 207–10

Lerner M M, Oriakhi C O 1997 In: Goldstein A N (ed.) *Handbook Nanophase Materials*. Dekker, New York

Liu, Y-J, DeGroot, D C, Schindler, J L, Kannewurf, C R, Kanatzidis, M G 1993 Stabilization of anilinium in vanadium (V) oxide xerogel and its post-intercalative polymerization to poly(aniline) in air. *J. Chem. Soc., Chem. Commun.* 593–6

Mann S 1993 Molecular tectonics in biomineralization and biomimetic materials chemistry. *Nature* **365**, 499–505

Mark J E, Wang S-B 1988 Reinforcement from alumina-type fillers precipitated into an elastomer. *Polym. Bull.* **20**, 443–8

Meldrum F C, Heywood B R, Mann S 1992 Magnetoferritin: *in vitro* synthesis of a novel magnetic protein. *Science* **257**, 522–3

Messersmith P B, Stupp S I 1992 Synthesis of nanocomposites: organoceramics. *J. Mater. Res.* **7** (9), 2599–611

Nandi M, Conklin J A, Salvati L Jr., Sen A 1991 Molecular level ceramic/polymer composites. 2. Synthesis of polymer-trapped silica and titania nanoclusters. *Chem. Mater.* **3**, 201–6

Nazar L F, Zhang Z, Zinkweg D 1992 Insertion of poly(*p*-phenylenevinylene) in layered MoO_3. *J. Am. Chem. Soc.* **114**, 6239–40

Riedel R, Dressler W 1996 Chemical formation of ceramics. *Ceram. Int.* **22**, 233–9

Wang S-B, Mark J E 1987 In-situ precipitation of reinforcing titania fillers. *Polym. Bull.* **17**, 271–7

Yanagida, S, Enokida, T, Shindo, A, Shiragami, T, Ogata, T, Fukumi, T, Sakaguchi, T, Mori, H, Sakata, T 1990 Preparation of quantized-CdS doped poly(methyl methacrylate) films. Optical and morphological properties. *Chem. Lett.* 1773–6

Yuan Y, Fendler J H, Cabasso I 1992 Photoelectron transfer mediated by size-quantized CdS particles in polymer-blend membranes. *Chem. Mater.* **4**, 312–8

Zhang S, Gonsalves K E 1994 Biomimetic Processing of calcium carbonate–chitosan composites In: Gonsalves K E, Chow G-M, Xiao T D, Cammarata R C (eds.) Molecularly Designed Ultrafine/Nanostructured Materials. Materials Research Society, Pittsburgh, PA, pp. 245–50

J. K. Premachandra
University of Cincinnati, Ohio, USA

Polymer–Silica Nanocomposites

Organic/inorganic nanocomposites are a class of composite materials that have been developed and have attracted significant interest since the early 1980s. Unlike the traditional composite materials which have macroscale domain sizes of micrometer and even millimeter scales, polymer–silica nanocomposites are mostly nanoscopic, with a typical physical constraint of 1–100 nm. There is also intimate contact between different phases in these materials. Therefore, they have the potential for providing unique combinations of properties that cannot be achieved by conventional methods.

Of organic/inorganic nanocomposites, polymer–silica nanocomposites draw the most attention, and their synthesis and properties are the focus of this article. The pioneering work in this area includes that from research groups directed by Schmidt, Mark, and Wilkes. There are several very informative reviews of organic/inorganic nanocomposites available (Chujo and Saegusa 1992, Novak 1993, Schmidt 1994, Sanchez and Ribot 1994, Mark *et al.* 1995, Mascia 1995, Mark 1996, Wilkes and Wen 1996, Wen and Wilkes 1996, Sharp 1998).

1. General Synthetic Considerations

Although polymer–silica nanocomposites can be prepared by a variety of approaches, most are obtained through the sol–gel approach: sol–gel reactions of silicon alkoxides and related derivatives. Considering the important role of the sol–gel reaction, it is difficult to understand the preparation of these nanocomposites without a basic knowledge of the sol–gel process. The sol–gel reaction of silicon alkoxides is a method for preparing inorganic silicon oxides under mild conditions (Brinker and Scherer 1990). It involves the simultaneous hydrolysis and condensation of silicon alkoxides to form a three-dimensional SiO_2 network. The reactions are carried out typically in an acidic or alkaline medium with either an acid or a base as catalyst. The general scheme is represented in Scheme 1.

Despite the fact that tetraethoxysilane (TEOS) results in considerable volume shrinkage, it is the most-used precursor to generate silica because of its moderate hydrolysis and condensation rate compared to the fast-reacting tetramethoxysilane (TMOS). The nature of the catalyst and the pH of the reaction have a pronounced effect on the hydrolysis and condensation reactions and also on the structure of the SiO_2 formed. The hydrolysis rate is high under an acidic environment relative to that of condensation, while base catalysis results in a higher condensation rate. Therefore, the presence of acid tends to produce more linear or polymeric molecules in the initial stage which leads to the formation of high-density, low-fractal-dimension structures, while a basic environment tends to produce more of a cluster growth leading to dense, colloidal, particulate structures. Once formed, the cosolvent(s) such as alcohol and small molecular weight by-products such as ethanol and water must be removed from the gelled SiO_2 network during the gel drying. The removal of these cosolvent(s) and by-products contributes to the high shrinkage that occurs during the classical sol–gel process. (For a discussion of possible remedies see *Polymer–Ceramic Nanocomposites: Control of Shrinkage**.)

Traditional sol–gel reactions of pure silicon alkoxides are very complicated, with many variables such as water-to-alkoxyl ratio, acid (base)-to-alkoxyl ratio, catalyst type, reaction medium, alkoxyl type, etc. The introduction of organic polymeric components adds additional variations. On the one hand this complexity makes the basic chemistry of the reactions even more difficult to understand, but on the other hand it opens the door for the development of many different synthetic approaches.

2. Synthetic Approaches

The unique low-temperature characteristic of sol–gel reactions provides opportunities to make polymer–silica hybrid materials at temperatures under which the polymeric species can survive. Numerous combinations of polymer–silica and different synthetic procedures have been tried in the attempt to create

Hydrolysis

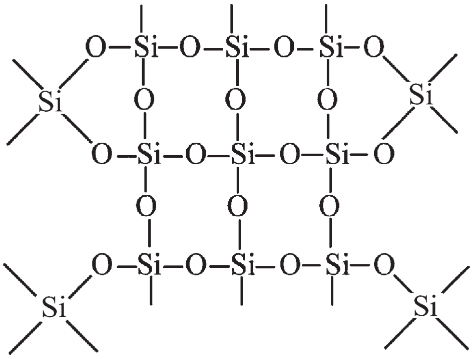
$$\begin{aligned}
Si(OR)_4 + H_2O &\rightleftharpoons (HO)Si(OR)_3 + ROH \\
(OH)Si(OR)_3 + H_2O &\rightleftharpoons (HO)_2Si(OR)_2 + ROH \\
(OH)_2Si(OR)_2 + H_2O &\rightleftharpoons (HO)_3Si(OR) + ROH \\
(OH)_3Si(OR) + H_2O &\rightleftharpoons Si(OH)_4 + ROH
\end{aligned} \qquad (1)$$

Condensation

$$\begin{aligned}
\equiv Si-OR + HO-Si\equiv &\rightleftharpoons \equiv Si-O-Si\equiv + ROH \\
\equiv Si-OH + HO-Si\equiv &\rightleftharpoons \equiv Si-O-Si\equiv + HOH
\end{aligned} \qquad (2)$$

Overall reaction

$$Si(OR)_4 \xrightarrow[H^+,\, OH^-]{H_2O,\ solvent}$$



(3)

Scheme 1

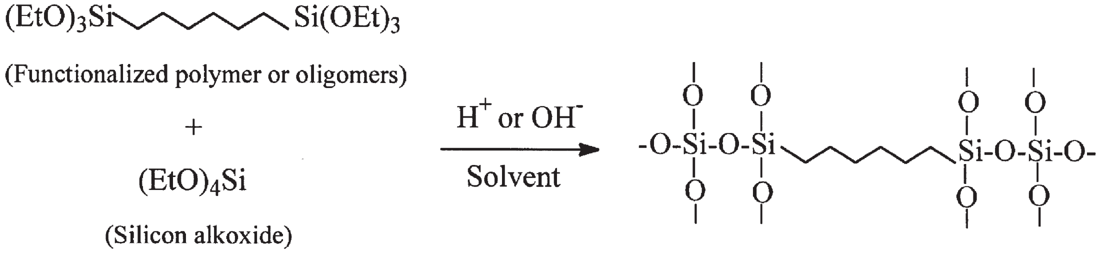
novel polymer–silica nanocomposites with vastly different properties. Therefore, a brief summary of different synthetic approaches is needed before the discussion of various polymer–silica nanocomposite systems.

2.1 Formation of Inorganic Silica Within a Swollen Polymer Network

An approach developed by Mark and Pan (1982) involves the *in situ* formation of inorganic silica within a swollen polymer network. In this approach, silica particles with a characteristic size of a few hundred angstroms can be generated within the polymer by first swelling a cross-linked polymeric host with a compatible solution containing a silicon alkoxide, usually TEOS, followed by the promotion of a base-catalyzed sol–gel reaction of the silicon alkoxide. This approach has been used successfully to reinforce many elastomers and plastics, such as poly(dimethylsiloxane) (PDMS), poly(methyl methacrylate) (PMMA), polyisobutylene, poly(vinyl acetate), polyurethane, etc. This approach has the most potential in the preparation of novel thin-film hybrid materials where the dispersion of water into the organic components is not a major concern. For a detailed review see *Elastomers, Ceramic-modified**.

2.2 Co-condensation of Functionalized Polymers with Silicon Alkoxides

Polymer–silica nanocomposites can be formed via co-condensation of functionalized polymers with silicon alkoxides in which chemical bonding is established between inorganic and organic phases. Scheme 2 shows this approach schematically with ethoxysilane



Scheme 2

functionalized polymer as an example. This approach was developed in the early 1980s (Wilkes *et al.* 1985) and has been used widely in the preparation of polymer–silica hybrid materials because of its ability to incorporate various organic polymers or oligomers with a silica network through chemical bonding. The typical terminal functional groups for polymers or oligomers include alkoxysilane, amino, hydroxyl, etc. Theoretically, any functional groups that can participate in the sol–gel reaction can be used to functionalize polymers or oligomers. However, a difference in the reactivity of the functional groups from that of the silicon alkoxide (typically TEOS) can result in poor incorporation of the organic component. Triethoxysilane is the most-used end group because its reactivity is comparable to that of TEOS and hence favors the formation of hybrid systems with better component dispersion and network development.

2.3 Simultaneous Formation of Silica and Organic Phases

Another approach to produce polymer–silica nanocomposites is through interpenetrating networks and simultaneous formation of both silica and organic polymeric phases. By using triethoxysilane, $R'_n(OR)_{3-n}$, as the precursor with R' being a polymerizable group such as an epoxy group, an organic network can be formed within the SiO_2 network by either photochemical or thermal curing of such groups (Schmidt *et al.* 1984). In this approach, the chemical bonding between inorganic and organic components already exists. Novak and Grubbs (1988) developed another similar method to form inorganic/organic simultaneous interpenetrating networks (SPINs), where both inorganic glass and polymer formation occur at the same time through a synchronous application of the aqueous ring-opening metathesis polymerization (aqueous ROMP) of cyclic alkenyl monomers and the hydrolysis and condensation of metal alkoxides. The successful preparation of polymer–silica nanocomposites using this approach requires matching the polymerization rates of the two systems as well as noninterference between the two reactions.

2.4 Infiltration of Oxide Gels with Organic Monomers or Mixing Polymers with Silicon Alkoxides

Polymer–silica nanocomposites can be obtained by either the infiltration of previously formed oxide gels with polymerizable organic monomers or the mixing of polymers with silicon alkoxides in a common solvent. In the first approach, the impregnation of porous silica gels with organics is followed by an *in situ* polymerization initiated by thermal or irradiation methods (Pope and Mackenzie 1989). In the second approach, polymers can be trapped within the oxide gel network if the hydrolysis and condensation of silicon alkoxides are carried out in the presence of polymers (Landry and Coltrain 1991). If no coupling agent is used, the interaction between organic and inorganic phases is typically physical in nature. The application of these approaches is apparently limited by the solubility of the polymers in the solvents.

2.5 Other Synthetic Strategies Using the Sol–Gel Process

Other synthetic strategies making use of the sol–gel process have also been developed. Given here are two examples of such approaches. A significant drawback associated with the sol–gel process is the removal of the cosolvent, excess water, and liberated alcohol. Large volume shrinkage during the drying process will occur due to the high interfacial area of the solid–liquid interface coupled with the reduction in chemical potential. By employing polymerizable monomers as the cosolvent such that all mixture components contribute either to the silica network or to the organic polymer to avoid large-scale shrinkage, "nonshrinking" sol–gel composite materials can be obtained (Ellsworth and Novak 1991). A different kind of inorganic/organic nanocomposite material has been developed by exploiting the unique intercalation and self-assembling characteristics of layered ceramics, particularly those based on the 2:1 layered silicates (Giannelis 1992). These 2:1 layered silicates consist of two-dimensional layers formed by sandwiching two SiO_2 tetrahedral sheets to an edge-shared octahedral sheet. Single chains of the polymer alternately stacked with the layers of the host can be obtained by intercalative polymerization of various monomers in the silicate galleries.

2.6 Synthesis Using Nanometer-size Silica Particles

Although the sol–gel approach is the dominant one, polymer–silica nanocomposites can also be obtained using nanometer-size silica particles. In this approach, silica particles are dispersed in either a polymer matrix or a monomer solution in which *in situ* polymerization of the monomer is carried out later. Silica particles are usually premodified so that a good dispersion of the particles in the matrix can be obtained. Polymer–silica nanocomposites such as polyamide 6–silica, PMMA–silica, nylon 11–silica, and polypyrrole–silica have been prepared using this approach.

3. Properties and Applications

Compared to the polymer–silica materials prepared by the conventional method (i.e., direct blending of silica with polymer), the polymer–silica nanocomposites discussed in this article often have superior properties and are expected to provide many different possibilities. The unique properties that these hybrid materials display arise mainly from three sources. (i) The reduction of the domain size to 100 nm or below (often much less, sometimes even at molecular level). A direct consequence of ultrasmall domain size (smaller than the wavelength of visible light) is optical clarity of the materials by virtual elimination of light scattering, although the refractive index of the two phases might be very different. As a result, polymer–silica nanocomposites find applications in many fields, such as optical devices, which are far beyond the scope of application of traditional composites. (ii) Generation of enormous interfacial areas due to the nanoscale domain size and intimate contact between the two phases. This can result in numerous covalent bonds or other interaction between the phases. Therefore, the formation of morphologies such as interpenetrating networks and even molecular-level mixing with no distinct phases that are essentially unknown for conventional composite materials become possible. (iii) The tremendous range of chemistry available to optimize specific properties and to produce materials with unique properties. It is possible to shift the morphology of these materials from that of a continuous flexible organic domain with separated inorganic phases to a continuous glassy inorganic phase with separated organic domains.

The sol–gel process makes possible the preparation of polymer–silica nanocomposites through numerous combinations of polymers and silicon alkoxides. The resulting hybrid materials could be vastly different in properties depending on the reaction conditions and the chemical structure of polymers and silicon alkoxides. Therefore, it is better to discuss the properties of individual polymer–silica systems (see Sect. 4). Generally speaking, polymer–silica nanocomposites usually show intermediate properties between those of organic polymers and inorganic silica glasses. Accordingly, they can be used to modify organic polymeric materials or to modify inorganic glassy silica. In terms of organic polymer modification, almost an infinite number of variations can be obtained while the flexibility is the focus in terms of inorganic glassy silica modification. Polymer–silica hybrid materials may also possess new properties that the original organic and inorganic components do not have because of their unique structural features, such as small domain size and intimate contact between the two phases. Besides the original role of enhancing mechanical properties, the applications for polymer–silica nanocomposites have expanded to many areas. Applications include reinforcement of elastomers and plastics, toughening of silica glasses, special coatings such as abrasion resistance coatings for plastics, electrical and nonlinear optical (NLO) materials for optical and optoelectronic devices, porous templates for catalysts and surfactants, and biomimic materials.

4. Examples of Polymer–Silica Nanocomposites

Many polymer–silica nanocomposites have been obtained successfully by different synthetic approaches. The resulting composites can vary from soft and flexible to brittle and hard materials depending on the chemical structure of the polymeric components and the overall organic-to-inorganic composition ratio. Some representative polymer systems, including some early classic systems, will now be presented individually. Other polymer–silica nanocomposite systems are discussed in Sanchez and Ribot (1994), Wilkes and Wen (1996), Wen and Wilkes (1996), Mark (1996), and Sharp (1998).

4.1 PDMS–Silica

One of the first successfully prepared polymer–silica nanocomposites was a PDMS–silica system developed first by Mark's group in the early 1980s. The use of PDMS is obvious considering the fact that the backbones of PDMS and silica possess the same nature of (Si–O–Si) chemical bonds which make the organic and inorganic components more compatible. In this approach, TEOS was the preferred precursor to generate the silica domain because of its good miscibility with PDMS and also its moderate hydrolysis and condensation rate compared to the fast-reacting TMOS. Three synthetic methods have been used to prepare polymer–silica nanocomposites. In the first approach, the already cross-linked PDMS elastomer is swollen with TEOS, and the sol–gel reactions are carried out within the PDMS matrix to generate *in situ* small, well-dispersed silica particles with a diameter around a few hundred angstroms. The silica particle size can be controlled by the ratio of PDMS to TEOS and the pH of the solution. In most cases base catalysts were used to promote the formation of dense silica particles. In the second technique, hydroxyl-terminated PDMS chains are blended with enough TEOS to both end-link them and provide silica by the hydrolysis reaction. Thus, curing and filler generation takes place simultaneously. In the third approach, TEOS is blended into PDMS having end groups (e.g., vinyls) that are unreactive under the hydrolysis conditions. Therefore, PDMS curing takes place after the generation of silica in the usual manner. It has been demonstrated that extremely uniform silica particle sizes can be obtained. Since the PDMS–silica nanocomposites prepared in this way can also be viewed as reinforced elastomers, they are the ideal candidates for studies of fundamental reinforcement principles. The concept of such "*in situ*" filler precipitation within elastomers is also of practical importance because these particles provide the same highly desirable reinforcing effects obtained by the traditional blending technique, yet avoid the time-consuming, energy-intensive process associated with it. For further details see *Elastomers, Ceramic-modified**.

PDMS–silica systems prepared with acid catalysts were also synthesized and studied by Wilkes and other researchers. Low molecular weight (500–1700 $\mathrm{g\,mol^{-1}}$) PDMS oligomers with silanol terminal groups and TEOS were used first as the reactants to form a chainlike organic/inorganic structure with hydrogen chloride (HCl) as catalyst. The studies were later expanded to the systems with high molecular weight PDMS, PDMS with higher functionality, and using other acid catalysts. Depending on the amount and molecular weight of PDMS, the final materials could be either flexible or brittle and all showed optical transparency indicating a phase (domain) size of smaller than the wavelength of visible light. The microstructure of these PDMS–silica hybrids was determined to be microphase-separated with PDMS- and silica-rich domains, using a variety of analytical techniques such as DSC, dynamic mechanical analysis (DMA), and small-angle x-ray scattering (SAXS). The structure and properties of PDMS–silica systems are found to be dependent mainly on the content of PDMS, the molecular weight of PDMS, and the content of the acid catalyst.

4.2 Acrylate Polymers–Silica

Acrylate polymers are another class of polymers that have attracted extensive research effort in the synthesis of polymer–silica nanocomposites. PMMA–silica nanocomposites have been synthesized using different approaches by different research groups. Schmidt synthesized PMMA–silica materials with improved mechanical properties through an *in situ* polymerization of methacrylate monomers, which were introduced by using monomers containing trialkoxysilyl functional groups, within the inorganic silica network. Through the infiltration of previously formed oxide gels with polymerizable organic monomers followed by an *in situ* polymerization initiated by thermal or irradiation methods, Pope and Mackenzie (1989) developed transparent optical PMMA–silica nanocomposites although there was no chemical bonding between the two phases without the use of a coupling agent. Extensive work has been done by Landry and Coltrain (1991) and Landry *et al.* (1992) toward synthesizing and studying PMMA–silica materials through first trapping PMMA within a silica gel and then carrying out the hydrolysis and condensation of a silicon alkoxide in the presence of PMMA. They found that there was no covalent bonding between organic and inorganic phases. The ability to produce well-dispersed nanocomposite materials was attributed to hydrogen bonding between the silanol groups in the inorganic phase and the carbonyl units in the PMMA.

Other acrylate–silica nanocomposites have also been produced and studied, including poly(methyl acrylate) (PMA), poly(butyl methacrylate) (PBMA), allyl methacrylate (AMA) and its copolymer with methyl methacrylate, and poly(methyl methacrylate-co-styrene). Based on their previous effort, Wei *et al.* (1998) have synthesized and studied a series of monolithic poly(methyl methacrylate-co-3-(trimethoxysilyl)propyl methacrylate)–silica hybrid materials through HCl-catalyzed sol–gel reactions of TEOS with polymethacrylates at various compositions. The structures of these hybrid materials were studied by atomic force microscopy (AFM) and other analytical techniques, and their applications as the filler for dental composite materials was also investigated. It was found that the polymethacrylate chains are uniformly distributed in, and covalently bonded to, the silica networks at a molecular level without macroscopic organic/inorganic phase separation.

4.3 Polyimides–Silica

Usually high-temperature polymers are very difficult to treat in the usual sol–gel technique. However, polyimides are synthesized from a precursor molecule, such as poly(amic acid), which is soluble in some organic solvents. This makes it possible to produce polyimide–silica hybrid materials through the sol–gel approach. Most of the research work has been aimed at improving the properties of polymers with respect to better thermal stability and mechanical strength. The synthetic procedure for the preparation of polyimide–silica nanocomposites involves mixing varying amounts of the precursor for the polyimide, typically poly(amic acid), with TEOS and water in a common solvent, followed by imidization and simultaneous network formation. Nandi *et al.* (1991) reported the preparation of molecular level polyimide–silica composites by using a so-called "site isolation" method in which TEOS is maintained in isolated pockets in the polymeric matrix by prebinding it to the carboxylic sites of the poly(amic acid) formed initially. The authors suggested that the metal alkoxide can be bonded chemically with the precursor poly(amic acid) through cohydrolysis of carboxylic groups and the metal alkoxide, although the formation of chemical bonding was not confirmed. Both transmission electron microscopy (TEM) and scanning electron microscopy (SEM) observations indicate the formation of a homogeneous dispersion of SiO_2 nanoclusters with a size of about 1.5 nm within the polyimide matrix. As a result, transparent and flexible polyimide-containing organic/inorganic films can be obtained in this way. Morikawa *et al.* (1992) also reported the preparation of polyimide–silica hybrid films by the hydrolysis and condensation of TEOS in a poly(amic acid) solution, followed by heating of the resulting film at high temperature. TEOS and water were added to the *N*,*N*′-dimethylacetamide (DMAc) solution of poly(amic acid). The reaction between TEOS and water resulted in a homogeneous mixture which was then cast onto a glass plate. The polyimide–silica hybrid film was obtained after heating to about 270 °C. Hybrid films having a silica content up to 70 wt.% and a silica size of approximately 5 μm can be obtained. In subsequent work, ethoxysilyl groups were introduced as the diamine monomer into the polyimide matrix to connect the silica particles with the matrix. By using methanol as the solvent, hybrid films containing up to 50 wt.% silica with much smaller size (about 0.2–0.08 μm in diameter) have been obtained. In contrast to the other preparation method given above, the hybrid films made in this way show a higher storage modulus, a reduction in the decrease of the glass transition temperature, and low tanδ maximum.

The preparation and properties of other polyimide–silica hybrid systems have been investigated by several other research groups. One focus of the research is to improve the compatibility between the two phases. Kioul and Mascia (1994) improved the compatibility of polyimide–silica ceramers by adding a small amount of γ-glycidyloxypropyltrimethoxysilane (GOTMS) as a coupling agent. The morphology of the hybrid films was dependent strongly on the molecular weight of the poly(amic acid), the alkoxysilane solution composition, the reaction time, and the type of coupling agent and catalyst. The tensile strength and ductility of the hybrid materials were found to increase with the addition of the low functionality alkoxysilane and GOTMS coupling agent. Using an organically substituted alkoxysilane, (aminophenyl)trimethoxysilane (APTMOS), to provide bonding sites between the polyimide and the silica-like phase, Wang *et al.* (1994) studied the morphology, thermal stability, and mechanical properties of polyimide–silica composites and found that small amounts of APTMOS can improve the modulus and strength of the resulting hybrid materials, and that the transparency also increased with an increase of APTMOS. In order to better control the structure, Hedrick *et al.* (1997) developed an alternative route to polyimide–silica hybrid materials based on poly(amic acid ethyl ester) precursors to the polyimide. Poly(amic acid ethyl ester) precursors are hydrolytically stable, allowing isolation, characterization, and copolymerization in a variety of solvents and cosolvents. The high cure temperature of poly(amic acid ethyl ester) precursors allows significant vitrification of the gel prior to imidization, providing a kinetic means of controlling the morphology. Compared to bulk silica, these polyimide–silica hybrids show enhanced resistance to cracking.

4.4 *Poly*(*Tetramethylene Oxide*) (*PTMO*)*–Silica*

Following their work on PDMS–silica nanocomposites, Wilkes and co-workers (Huang 1990) turned their attention to PTMO–silica systems and conducted systematic investigations on the preparation, structure, and properties of PTMO–silica hybrid materials. PDMS–silica nanocomposites did not display high tensile strength and elongation, which was attributed to the inherently poor mechanical properties of pure PDMS and the network defects such as remaining loose or dangling PDMS chain ends and the molecular weight redistribution of PDMS chains. Therefore, PTMS–silica, which can offer much better mechanical properties, was made to overcome these problems. However, the lack of the siloxane backbone to make the organic polymer and inorganic silica compatible was a concern. To improve the compatibility of PTMO with silica, isocyanatopropyltriethoxysilane (instead of hydroxyl) functionalized PTMO oligomers were used. Depending on the reaction conditions and composition ratio of PTMO to TEOS, these resulting hybrid materials can vary from soft and flexible to brittle and hard materials and still maintain optical transparency. Compared to PDMS–silica systems, the acid-catalyzed PTMO-containing hybrid materials generally provide an enhanced mechanical response (30 MPa in tensile strength and 100% in elongation compared to 6 MPa and 20%) for an equivalent oligomer content due to higher functionality (6 compared to 2) and the nature of PTMO oligomers.

The structure and morphology of PTMO–silica nanocomposites were investigated extensively using SAXS and dynamic mechanical spectroscopy (DMS) techniques. The high transparency, given the large difference in refractive index between the two components, clearly demonstrated that no macroscale phase separation occurred. However, SAXS results showed that they tend to undergo some microphase separation similar to segmented or block polymers. The observed correlation distance is of the order of about 10 nm for the PTMO–silica systems. The high angular tail of the scattering curve also suggested that some level of mixing of the inorganic phase occurs within the oligomeric phase. This partial phase mixing was also supported by DMS data.

Systematic work has also been done on the variables of oligomer molecular weight and functionality, ratio of PTMO to TEOS, acid concentration and type, processing temperatures and methods, gel aging, and solvent type by using characterization techniques such as SAXS, DMS, IR spectroscopy, NMR, and mechanical testing. Besides mineral and low molecular weight organic acids the use of the polymeric acid catalyst poly(styrene sulfonic acid) (PSS), which is readily available in various molecular weights and degrees of sulfonation, has also been investigated. An advantage of a polymeric acid catalyst is that it can be used as a rheological modifier to enhance processability yet minimize fugitive acid moieties that are prominent in a low molecular weight catalyst. SAXS studies also showed that when using a *N*,*N*-dimethylformamide (DMF)–IPA cosolvent system as the reaction medium, instead of using pure IPA as solvent, the SAXS profiles became much sharper as the DMF content increased, indicating a narrower distribution of either domain sizes or interdomain spacings or both. This was believed to be attributable to the less acidic nature of the mixture when DMF was used as cosolvent because it promoted less mixing between the two components.

Other studies in Wilkes' program include the chemical stability of PTMO–silica hybrid materials in neutral and alkaline aqueous solutions, the aging effect, and the microwave curing process. It was found that because of the solubility of silicon oxide in highly basic solutions a better-defined two-phase microstructure can be obtained if PTMO–silica hybrids were aged in 1 M NaOH for 24 h. If the films cast at ambient temperature are later heated to a higher temperature (~100 °C), the extent of the reaction can be increased. This so-called "aging" can also take place in the time period between gelation and vitrification of the "hard" regions and this aging effect can be easily interpreted by using the concept of the time–temperature transformation (TTT) cure diagram established by Gillham. Microwave curing processing has also been used to promote the sol–gel reaction of PTMO–silica systems. Microwave processing has the ability to promote potentially a higher rate of reaction due to specific interactions with the dipoles of the appropriate functionalized reacting groups. It was found that gelation and near final properties of the PTMO–silica hybrid systems can be achieved in a matter of minutes. The higher intensity and flat "tail region" from the SAXS curve of microwave-processed PTMO–silica hybrid networks also indicated a higher extent of reaction and higher level of phase separation compared to ambient-temperature casting or even conventional thermal (convection) cured systems.

4.5 *Polyoxazoline (POZO)–Silica*

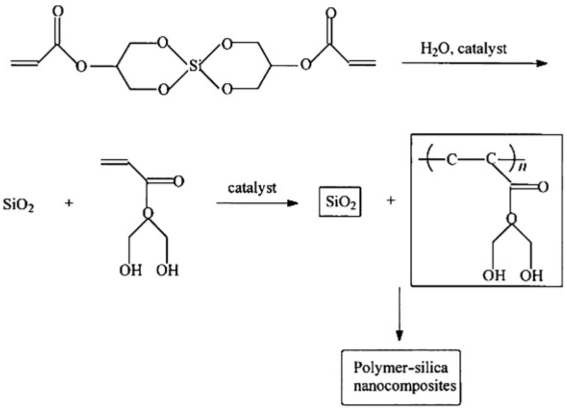
By using a comparable synthetic approach, Chujo *et al.* (1991) synthesized POZO–silica nanocomposites using triethoxysilane-terminated POZO with a molecular weight of 500–2000 $\mathrm{g\,mol^{-1}}$ for the acid-catalyzed cohydrolysis and condensation with TEOS. The resulting materials are homogeneous, transparent, and glassy inorganic/organic nanocomposites with chemical bonding between the two phases. Similar POZO–silica hybrid materials which have no chemical bonding between the organic and inorganic phases have also been prepared by using the hydrogen bond interactions between the silanol group and the POZO amide carbonyl group. The latter material has been used as the precursor for porous silica glass of controllable pore size and very large surface area by calcination of the materials at temperatures between 600 and 800 °C. Poly(ethyl oxazoline) and poly(methyl oxazoline) have also been dispersed into a silica matrix to make POZO–silica hybrid materials (David and Scherer 1995). The authors claim that the dispersion is at the molecular level because only a single phase was observed as evidenced by SAXS and TEM results.

4.6 *Poly(Ether Ketone) (PEK)–Silica*

PEK–silica hybrids were prepared with a similar synthetic approach to that of the PTMO–silica systems, using triethoxysilane-functionalized PEK oligomers ($M_n = 3900\ \mathrm{g\,mol^{-1}}$) and TEOS (Noell *et al.* 1990). PEK is an amorphous thermoplastic material with good mechanical properties, thermal stability, and a relatively high glass transition temperature compared to PDMS or PTMO systems. Although this material is transparent, SAXS experiments confirmed the presence of local microphase separation, as is the case with other hybrid materials. However, the high glass transition temperature of PEK can limit further the extent of cure caused by vitrification as gelation and network buildup occur. Again, thermal postcure treatments at elevated temperature can promote further network development.

4.7 *Nonshrinking Polymer–Silica*

It is well known that in the sol–gel process there is significant gel shrinkage and cracking during drying



Scheme 3

due to the removal of excess water, cosolvent(s), and the liberated alcohol. This shrinkage is caused mainly by the large capillary forces generated in the nanoporous structure. As a consequence, films with a thickness larger than 1 μm are difficult to dry and usually crack extensively. One way to alleviate the shrinkage problem is to incorporate organic polymers into the inorganic silica network. However, shrinkage still remains a problem for the preparation of polymer–silica nanocomposites. Novak and Davis (1991) and Ellsworth and Novak (1991) developed a way to synthesize so-called "nonshrinking" polymer–silica hybrid materials. They synthesized a series of tetraalkyl orthosilicate derivatives possessing polymerizable alkoxide groups. These siloxane derivatives liberate four equivalents of polymerizable alcohol during the sol–gel process. By employing polymerizable monomers as cosolvent, all components contribute either to the silica network or to the organic polymer. Therefore, gel drying is unnecessary and large-scale shrinkage is avoided. Scheme 3 shows a typical example of this approach. The mechanical properties of these materials can be improved greatly through the incorporation of an organic phase. The glass content of these nonshrinking hybrid materials can be increased further by using soluble poly(silicic acid ester) oligomers possessing the same polymerizable alkoxides as the staring materials.

4.8 *Poly(ε-Caprolactone) (PCL)–Silica*

This is a typical example of a polyester–silica system. Tian *et al.* (1997, 1999) reported the synthesis and characterization, and particularly the property–structure–morphology characteristics, of biodegradable and biocompatible PCL–silica hybrid materials. PCL is well known for biocompatibility, permeability, and biodegradability. Transparent hybrid materials have been prepared from both linear and three-arm star-shaped triethoxysilane-, hydroxyl-,or vinyl-terminated PCL and TEOS by the acid-catalyzed sol–gel process. The organic PCL components and inorganic silica components can be associated either by covalent bonding (in the case of triethoxysilane-terminated PCL) or by hydrogen bonding between the carbonyl groups of PCL and the residual OH groups on the silica. The incorporation of PCL into the silica network, and consequently the morphology and structure of the resulting PCL–silica materials, has been investigated by TEM, SAXS, DMA, AFM, SEM-EDAX, IR analysis, and solid-state NMR, and was found to be dependent on the PCL content and molecular weight, the reactivity of the PCL end-groups and their number per chain, and the acid and water content. For example, lower PCL content, higher PCL molecular weight, higher acidity, more functional groups per chain, and triethoxysilane end groups are favorable to PCL incorporation. The analysis of the Porod region of the SAXS profiles show that the PCL–silica interface is mass fractal, in agreement with a co-continuous two-phase morphology. A more open structure can be formed at lower acid content and higher water content. Similar to PTMO–silica and PEK–silica systems, the extent of the sol–gel reaction is limited by the early vitrification of the reactive system, and thermal post-treatment can overcome these diffusional and/or kinetic limitations. *In vitro* cell culture and biodegradation tests show that PCL–silica systems are biomaterials with biodegradable and biocompatible properties. Another application of these materials is that they can be used to produce porous silicas by pyrolysis to remove the organic components. The silica microporosity can be tailored by the content, molecular weight, and molecular weight distribution of PCL as well as by the reactivity and number of the PCL end groups.

4.9 *Partially Deacetylated Chitin (CHI)–Silica*

CHI is a biopolymer structurally similar to cellulose, but containing predominantly amino groups at the C-2 position. CHI–silica hybrid materials were made by a controlled acid-catalyzed hydrolysis of TEOS via the sol–gel process in the presence of CHI (Retuert *et al.* 1997). The resulting hybrids were homogeneous, transparent, glassy materials with a wide composition range. The hydrogen bonding between the amino functional groups of CHI and the silanol groups on the silica could be attributed to some degree to the compatibility between CHI and silica. It was also suggested that the homogeneity of the hybrids could be due to the dispersion of CHI at the molecular level in the silica network.

4.10 *Fluoropolymers–Silica*

Generally speaking, fluoropolymers are not good candidates for preparing polymer–silica nanocomposites because most of them only dissolve in perfluorinated solvents that are profoundly incompatible with water. Work has been done to conduct the sol–gel reactions in such highly inhospitable solvents by using polyfluoroalkoxysilanes with trifluoroacetic acid as the gelling agent (Sharp 1998). For example, *in situ* generation of silica by the sol–gel process was carried out in the presence of the soluble fluoropolymer Teflon AF which is a copolymer of tetrafluoroethylene and bis-2,2-(trifluoromethyl)-4,5-difluoro-1,3-dioxole. Optically clear coatings can be obtained after casting the solution on glass or metal substrates. These hybrid materials show improved adhesion to glass and better abrasion resistance than the pure fluoropolymer counterpart. SAXS and TEM results showed that the largest particles observed in a hybrid containing 17 wt.% silicate in Teflon AF were about 15 nm. Therefore, the extremely high transparency of the fluoropolymer was not compromised by the presence of moderate amounts of silicate.

Bibliography

Brinker C J, Scherer G W 1990 *Sol–Gel Science: The Physics and Chemistry of Sol–Gel Processing*. Academic Press, New York

Chujo Y, Ihara E, Kure S, Suzuki K, Saegusa T 1991 Block copolymer of 2-methyl-2-oxazoline with silica gel. An organic–inorganic hybrid polymer. *Makromol. Chem. Macromol. Symp.* **42/43**, 303

Chujo Y, Saegusa T 1992 Organic polymer hybrids with silica gel formed by means of the sol–gel method. *Adv. Polym. Sci.* **100**, 11–29

David I A, Scherer G W 1995 An organic/inorganic single-phase composite. *Chem. Mater.* **7**, 1957–67

Ellsworth M W, Novak B M 1991 Mutually interpenetrating inorganic–organic networks. New routes into nonshrinking sol–gel composite materials. *J. Am. Chem. Soc.* **113**, 2756–8

Giannelis E P 1992 A new strategy for synthesizing polymer–ceramic nanocomposites. *JOM* **44**, 28

Hedrick J L, Cha H-J, Miller R D, Yoon D Y, Brown H R, Srinivasan S, Pietro R D 1997 Polymeric organic–inorganic hybrid nanocomposites: preparation of polyimide-modified poly(silsesquioxane) using functionalized poly(amic acid alkyl ester) precursors. *Macromolecules* **30**, 8512–5

Huang H-H, Glaser R H, Wilkes G L 1988 New hybrid materials incorporating poly(tetramethylene oxide) into tetraethoxysilane-based sol–gel glasses. In: Zeldin M, Wynne K J, Allcock H R (eds.) *Inorganic and Organometallic Polymers: Macromolecules Containing Silicon, Phosphorus, and Other Inorganic Elements*, ACS Symp. Ser. 360. American Chemical Society, Washington DC, Chap. 29. p. 354

Kioul A, Mascia L 1994 Compatibility of polyimide–silicate ceramers induced by alkoxysilane silane coupling agents. *J. Non-Cryst. Solids.* **175**, 169

Landry C J T, Coltrain B K 1991 Organic–inorganic composites prepared via the sol–gel method: some insights into morphology control and physical properties. *Polym. Prepr.* **32** (3), 514–5

Landry C J T, Coltrain B K, Wesson J A, Zumbulyadis N, Lippert J L 1992 In situ polymerization of tetraethoxysilane in polymers: chemical nature of the interactions. *Polymer* **33**, 1496–506

Mark J E 1996 The sol–gel route to inorganic–organic composites. *Heterog. Chem. Rev.* **3** (4), 307–26

Mark J E, Lee C Y-C, Blanconi P A (eds.) 1995 *Hybrid Organic–Inorganic Composites*, ACS Ser. 585. American Chemical Society, Washington DC

Mark J E, Pan S-J 1982 Reinforcement of poly(dimethylsiloxane) networks by in-situ precipitation of silica: a new method for preparation of filled elastomers. *Makromol. Chem. Rapid Commun.* **3**, 681–5

Mascia L 1995 Developments in organic–inorganic polymeric hybrids: creamers. *Trends Polym. Sci.* **3**, 61

Morikawa A, Iyoku Y, Kakimoto M, Imai Y 1992 Preparation of new polyimide–silica hybrid materials via the sol–gel process. *J. Mater. Chem.* **2**, 679

Nandi M, Conklin J A, Salvati J Jr, Sen A 1991 Molecular level ceramic/polymer composites: 2. Synthesis of polymer-trapped silica and titania nanoclusters. *Chem. Mater.* **3**, 201

Noell J L W, Wilkes G L, Mohanty D K, McGrath J E 1990 The preparation and characterization of new polyether ketone–tetraethylorthosilicate hybrid glasses by the sol–gel method. *J. Appl. Polym. Sci.* **40**, 1177–94

Novak B M 1993 Hybrid nanocomposite materials—between inorganic glasses and organic polymers. *Adv. Mater.* **5**, 422

Novak B M, Davis C 1991 "Inverse" organic–inorganic composite materials: 2. Free-radical routes into nonshrinking sol–gel composites. *Macromolecules* **24**, 5481

Novak B M, Grubbs R H 1988 Catalytic organometallic chemistry in water: the aqueous ring-opening methathesis polymerization of 7-oxanorbornene derivatives. *J. Am. Chem. Soc.* **110**, 7542–3

Pope E J A, Mackenzie J D 1989 Transparent silica gel–PMMA composites. *J. Mater. Res.* **4**, 1017

Retuert J, Nuñez A, Martinez F, Yazdani-Pedram M 1997 Synthesis of polymeric organic–inorganic hybrid materials. Partially deacetylated chitin–silica hybrid. *Makromol. Chem. Rapid Commun.* **18**, 163–7

Sanchez C, Ribot F 1994 Design of hybrid organic–inorganic materials synthesized via sol–gel chemistry. *New J. Chem.* **18**, 1007

Schmidt H 1994 Inorganic–organic composites by sol–gel techniques. *J. Sol–Gel Sci. Tech.* **1**, 217

Schmidt H, Scholze H, Kaiser A 1984 Principles of hydrolysis and condensation reactions of alkoxysilanes. *J. Non-Cryst. Solids* **63**, 1–11

Sharp K G 1998 Inorganic/organic hybrid materials. *Adv. Mater.* **10**, 1243

Tian D, Blacher S, Jerome R 1999 Biodegradable and biocompatible inorganic–organic hybrid materials: 4. Effect of acid content and water content on the incorporation of aliphatic polyesters into silica by the sol–gel process. *Polymer* **40**, 951–7

Tian D, Dubois P H, Jérôme R 1997 Biodegradable and biocompatible inorganic–organic hybrid materials: 1. Synthesis and characterization. *J. Polym. Sci. Part A.* **35**, 2295–309

Wang S, Ahmad Z, Mark J E 1994 Polyimide–silica hybrid materials modified by incorporation of an organically substituted alkoxysilane. *Chem. Mater.* **6**, 943–6

Wei Y, Jin D, Wei G, Yang D, Xu J 1998 Novel organic–inorganic chemical hybrid fillers for dental composite materials. *J. Appl. Polym. Sci.* **70**, 1689–99

Wen J, Wilkes G L 1996 Organic/inorganic hybrid network materials by the sol–gel approach. *Chem. Mater.* **8**, 1667–81

Wilkes G L, Orler B, Huang H 1985 "Ceramers": hybrid materials incorporating polymeric/oligomeric species into inorganic glasses utilizing a sol–gel approach. *Polym. Prep.* **26** (2), 300–1

Wilkes G L, Wen J 1996 Inorganic/organic composites. In: Salamone J C (ed.) *The Polymeric Materials Encyclopedia: Synthesis, Properties and Applications*. CRC Press, Boca Raton, FL, Vol. 6, pp. 4782–92

J. Wen
Cabot Corporation, Tuscola, Illinois, USA

Polymers in Constrained Ceramic Environments

There is a great deal of interest in the preparation and characterization of organic–inorganic nanocomposites, particularly in those materials in which an organic phase is ceramic constrained in such a way as to produce novel structures, frequently at the molecular level. Such ceramic-constrained environments include zeolites, clay galleries, and mesoporous silica. In some cases, a polymer is directly introduced into a constraining structure (clay galleries), while in others a constrained monomer is polymerized subsequently, for example in pores, to obtain the final composite material. This review focuses on some specific aspects of the morphology of organic–inorganic nanocomposites and how the confinement affects the properties of the polymer.

1. Simple Mixture of a Polymer and a Zeolite

In the simplest study of this type, Al-Ghamdi and Mark (1988) studied reinforcement of poly(dimethylsiloxane) (PDMS) by two zeolites having different pore sizes. The zeolites were simply blended into hydroxyl-terminated chains of PDMS, which were subsequently end-linked with tetraethoxysilane to form an elastomeric network. Stress–strain isotherms measured in elongation showed that both types of zeolites increased the ultimate properties of the elastomer and were thus effective reinforcing fillers, the one having the larger pore size giving a larger increase in the ultimate strength of the elastomers.

Further characterization of the mechanical properties and structure of such zeolite-reinforced PDMS elastomers by Wen and Mark (1994) also utilized small-angle neutron scattering (SANS) and transmission electron microscopy (TEM). The neutron-scattering profiles of the pure and zeolite-filled PDMS networks were identical, which indicates negligible penetration of the polymer into the zeolite pores. The TEM pictures showed that the zeolite with the larger pore size (zeolite 13X) had a smaller particle size, and this is probably the origin of its superior reinforcing ability.

2. Polymerizing Monomer within Zeolite Cavities

A different approach was initiated by Frisch and Xue (1995) to introduce polymer into the internal pores of the zeolite 13X. As will be seen in the following section, the approach is very similar to that used by Wu and Bein (1994a,b) to form molecular-scale wires of polyaniline within nanometer-sized channels.

In the first application of this approach, Frisch and Xue added (liquid) monomeric styrene in various ratios to a weighed amount of zeolite. A desired amount of benzoyl peroxide was then added to the mixture for subsequent generation of free radicals by thermolysis, producing polymerization of the styrene. The polymerization was carried out at 70 °C, either directly in this form or in the presence of the cross-linker ethylene glycol dimethacrylate (EGDM; 0.017 mol/mol of monomer). In the absence of the cross-linker, the linear PS chains and the zeolite were considered to form a hybrid inorganic/organic "pseudo-interpenetrating polymer network" (PIPN) or "semi-interpenetrating network," while the cross-linked polystyrene (PS) and zeolite composite was considered to have formed a full interpenetrating polymer network (IPN).

Subsequently, Frisch *et al.* (1996a, c) were also able to form both PIPNs and IPNs from poly(ethylene acrylate) (PEA) and zeolite 13X. At room temperature the PEA is a rubbery polymer while PS is glassy. In this case, the polymerization was carried out at 40 °C using liquid ethyl acrylate monomer in the zeolite mixtures, with benzoyl peroxide initiator, either in the absence or presence of EGDM cross-linker.

These studies provided substantial confirmation that the polymer grows within the internal pores. Specifically:

(i) Removal of about half of the linear polymer by solvent extraction (determined gravimetrically) left individual zeolite crystals containing the remaining polymer. The SEM micrographs showed that there was essentially no polymer between or on the zeolite surface (which was present before the extraction), and resembled micrographs of the pure zeolite. All linear polymer could be removed on extended solvent extraction (Soxhlet extraction for one week in suitable solvents).

(ii) Solid-state ^{13}C nuclear magnetic resonance (NMR) showed significant line-width broadening of the quaternary aromatic, methine, and methylene carbons in the linear PS. The effect was strongest in the partially solvent-extracted samples. This is consistent with the speculation that the "internal" PS chains in the pores of the zeolite are more confined by their environment than in the pure PS (without any zeolite).

(iii) The small-angle x-ray scattering (SAXS) curve of the PIPN of the PEA/zeolite exhibited two maxima, as shown in Fig. 1. The first rather sharp peak was the same as in the pure zeolite 13X and corresponds to a Bragg's law distance of 14.3 Å (the periodic interpore distance in zeolite 13X). The second (weaker and broader) peak disappears completely in the scattering curve for the fully extracted PIPN (containing no residual PEA polymer). This weaker peak has a maximum which corresponds to the doubled distance (∼28.6 Å) which is interpreted to be due to the presence of PEA-filled pores in an arrangement showing some type of imperfect regularity.

(iv) Of greatest relevance were the differential scanning calorimetry (DSC) curves for IPNs and PIPNs of PS/zeolite or PEA/zeolite containing 50–60 wt.% polymer (where the zeolite crystals were present during the polymerizations, with some of them being threaded in a matrix of cross-linked or linear polymer). These materials exhibited a glass transition temperature T_g, which was approximately the same as that of the corresponding pure polymer. Partially extracted samples of the PIPNs in which essentially all the interparticle and surface polymer has been removed exhibited no T_g at all, despite gravimetric weight measurements indicating that a considerable amount of PS (26.6 wt.% PS) was still present. Presumably, the internal linear polymer chains in the pores of the zeolite 13X adopt a conformation more extended in one dimension and thus exhibit no bulk T_g. In qualitative terms, the constraints prevent the long-range motions that would otherwise occur at the normal glass transition. The same results were found with partially extracted PEA in zeolite 13X. The root mean square radii of gyration of polystyrene and polyethylacrylate of viscosity average molecular weight of 3.0×10^5 for PS and 4.2×10^5 for PEA which were used are 71 Å and 99 Å, respectively. In the zeolite 13X the pores have a diameter of 10 Å, considerably smaller than the root mean square radii of gyration of these polymers.

3. Polymers Constrained in Clay Galleries

A polymer–clay hybrid is defined as a molecular composite of polymer and silicate layers of clay minerals, such as montmorillonite, saponite, hectrite, and synthetic mica. Two main methods of synthesis have been proposed. One is the *in situ* polymerization of metal alkoxides in organic matrices via the sol–gel process. The second is the utilization of intercalation reactions of clay minerals. This article reviews the research in the latter field in which polymers have been introduced into suitably modified interlayer spaces (galleries) of clays.

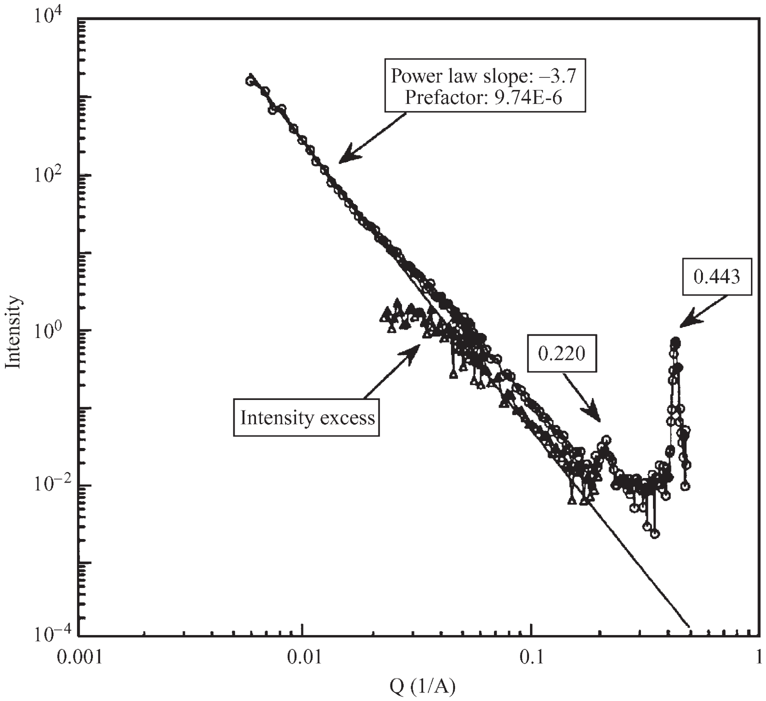


Figure 1
Excess scattering intensities for a poly(ethyl acrylate)/zeolite pseudointerpenetrating polymer network (PIPN).

In general, two types of hybrids are possible: intercalated, in which a single, extended polymer chain is intercalated between the host layers resulting in a well ordered multilayer with alternating polymer/inorganic host layers and repeat distance of a few nanometers; and delaminated, in which the silicate layers (1 nm thick) are exfoliated and dispersed in a continuous polymer matrix.

In an early study in this area, Vaia *et al.* (1993) reported on the properties of the intercalated polymers by concentrating on the polystyrene–organosilicate hybrid. The gallery height increased by 0.7 nm in the hybrid. The intercalated polystyrene chains adopted a collapsed, two-dimensional conformation and did not show the characteristic bulk glass transition. This behavior is qualitatively different from that exhibited by the bulk polymer and is attributed to the confinement of the polymer chains between the host layers.

Kojima *et al.* (1993b,c) synthesized nylon 6–clay hybrid, in which layered clay minerals are uniformly dispersed in nylon 6 matrix. The features of this polymer are high strength, high modulus, high distortion temperature, and low gas permeability. It was suggested that the ionic interaction between the host and the polymer was one of the reasons why these hybrid materials had superior mechanical properties.

Kojima *et al.* (1993a) also synthesized rubber–clay hybrid (RCH). The mechanical properties of RCH were superior to those of vulcanized acrylonitrile butadiene rubber. The permeabilities of hydrogen and water vapor in RCH were measured and it was found that the RCH has an improved gas barrier property. The same result was found by Yano *et al.* (1993) in synthesizing the polyimide–clay hybrid. Furthermore, this hybrid also had a low thermal expansion coefficient.

Okada and Usuki (1995) prepared several polymer–clay hybrids, in which the polymers are nylon 6, nitrile rubber, and others. These, when molded, have excellent mechanical, thermal, and chemical properties as compared with the unfilled polymers and/or conventional composites. The mechanism of reinforcement was explained using the results of ^{15}N CP-MAS and proton pulsed NMR studies. The results revealed that all polymers in the interlayer space of minerals are partially ionized. It is suggested that the chemical bond between the inorganic silicate and the organic polymers at the molecular level causes an improvement in the macroscopic mechanical properties of materials.

In the work of Krishnamoorti *et al.* (1996), the static and dynamic properties of polymer-layered silicate nanocomposites are discussed in the context of polymers in confined media. Despite the topological constraints imposed by the host lattice, mass transport of the polymer into the silicate layers (at least in the case of essentially non-polar polystyrene) appears to be unhindered and exhibits mobility similar to that of the pure polymer. However, both the local and global dynamics of the polymer in the hybrids are dramatically different from those in the bulk. On a local scale, intercalated polymer chains exhibit higher flexibility along their backbone along with a marked suppression (or even absence) of cooperative dynamics typically associated with the glass transition. On a global scale, relaxation of polymer chains either tethered to or in close proximity (<1 nm as in intercalated hybrids) to the host surface are dramatically altered and parallel those of other intrinsically anisotropic materials such as block copolymers and liquid crystals.

Lee and Jang (1996) prepared a polymethylmethacrylate (PMMA)–clay hybrid, and the differential scanning calorimetry (DSC) traces corroborate the confinement of the polymer in the inorganic layer by exhibiting no observable transition in the thermogram. Both the thermal stability and tensile properties of the products appear to be substantially enhanced.

Later, the relaxation behavior of poly(ethylene oxide) (PEO), intercalated in montmorillonite, was studied by Vaia *et al.* (1997). The motions of the intercalated PEO chains are inherently noncooperative relative to the cooperative T_g motions in the amorphous portion of the bulk polymer. This is presumably due to the strong confining effect of the silicate layers on the relaxation of the intercalated polymer. Additional details are given in *Clay-based Polymer Nanocomposites** in this Encyclopedia.

The layered silicate–epoxy nanocomposites have also been investigated by Messersmith and Giannelis (1994). The nanocomposite exhibits a broadened T_g at slightly higher temperature than the unmodified epoxy. Furthermore, the dynamic storage modulus of the nanocomposite containing 4 vol.% silicate was approximately 58% higher in the glassy region and 450% higher in the rubbery plateau region compared to the unmodified epoxy.

4. Polymerization of Monomer within Porous Silica

After the study of the zeolite/polymer composites, Frisch *et al.* (1996b) also investigated the polystyrene (PS) constrained in two porous silicas, one of which was an amorphous silica composed of silica gel nanotubes with pore sizes of 200–8000 Å, and the other a novel crystalline mesoporous hexagonal silica with uniform holes of 28 Å in diameter.

Because of the large ratio of the silica pore diameter to the PS root mean square radius of gyration, the interpenetrating network (IPN) and pseudointerpenetrating network (PIPN) of the nanotube silica material displayed normal glass transitions, which differed little from the pure cross-linked PS and linear PS transitions. The small difference can be attributed to molecular weight differences, with lower molecular weights corresponding to lower glass transition temperatures. The extracted PS PIPN material in the nanotubes retained the bulk glass transition, unlike the PS confined in the smaller diameter zeolite 13X in which it disappeared. However, the pore size of the nanosilica was one to three orders of magnitude larger, with a diameter of 200–8000 Å compared to the zeolite 13X diameter of 10 Å. The morphology of the chains in the nanotubes may be similar to bulk PS and thus exhibit a bulk T_g.

The PS PIPN material made with the mesoporous silica was very similar in physical appearance to the nanosilica PIPN. The unextracted material had a normal glass transition curve and a T_g close to the bulk PS T_g. However, upon extraction of the polymer in the PIPN composite as the wt.% of organic polymer to inorganic material dropped from 70:30 to 22:78, when the external PS was removed there was no bulk glass transition observed in the temperature range 25–160 °C. This suggests that no bulk glass transition exists for linear PS chains (root mean square radius of gyration of about 100 Å) confined to pores of less than 30 Å.

Further work on the mechanical properties of films of such hybrid inorganic/organic IPNs is in progress (Maaref *et al.* 1999). The authors' group have investigated the conductive polyaniline complex/zeolite hybrids at high polyaniline contents which exhibit an apppreciable antiferromagnetic component as well as paramagnetic component to their electron paramagnetic resonance spectra (Frisch *et al.* 2001).

5. *Concluding Remarks*

The effect on the mechanical properties of the polymer (e.g., reinforcement) from forming a polymer–ceramic mixture can be quite variable and depend strongly on the particle size, etc. The mechanical properties of polymers confined to clay galleries are generally superior. This may be due, at least in part, to the fact that in certain clay systems there are ionic forces between the clay gallery and the polymer. When the distances between the walls of the clay gallery are sufficiently small then the polymer morphology in this constrained structure changes and no bulk T_g is observed. In general, when polymers are confined in pores whose diameter is smaller than the root mean square radius of gyration of the polymers no bulk T_g is observed. When the diameter of the pore is larger than the root mean square radius of gyration, as in the silica nanotubes, a bulklike T_g is observed. Clearly confinement in these cases also changes the polymer morphology with consequences for the glass transition.

Bibliography

Al-ghamdi A M S, Mark J E 1988 Zeolites as reinforcing fillers in an elastomer. *Polym. Bull.* **20**, 537–42

Frisch H L, Maaref S, Xue Y, Beaucage G, Pu Z, Mark J E 1996a Interpenetrating and pseudo-interpenetrating polymer networks of polyethylacrylate and zeolite 13X. *J. Polym. Sci., Polym. Chem. Ed.* **34**, 673–77

Frisch H L, Mark J E 1996 Nanocomposites prepared by threading polymer chains through zeolites, mesoporous silica, or silica nanotubes. *Chem. Mater.* **8**, 1735–8

Frisch H L, Song H, Ma J, Rajailovich M, Zhu S, Gambino R, Yang N L, Yan X 2001 An EPR study of polyaniline salt–zeolite nanocanposite. To be submitted

Frisch H L, West J M, Goltner C G, Attard G S 1996b Pseudo IPNs and IPNs of two porous silicas and polystyrene. *J. Polym. Sci., Polym. Chem. Ed.* **34**, 1823–6

Frisch H L, Xue Y 1995 Hybrid inorganic/organic interpenetrating polymer networks based on zeolite 13X and polystyrene. *J. Polym. Sci., Polym. Chem. Ed.* **33**, 1979–85

Frisch H L, Xue Y, Maaref S 1996c Pseudo interpenetrating polymer networks and interpenetrating polymer networks of zeolite 13X and polystyrene and poly(ethyl acrylate). *Macromol. Symp.* **106**, 147–66

Kojima Y, Fukumori K, Usuki A, Okada A, Kurauchi T 1993a Gas permeabilities in rubber-clay hybrid. *J. Mater. Sci. Lett.* **12**, 889–90

Kojima Y, Usuki A, Kawasumi M, Okada A, Kurauchi T, Kamigaito O 1993b Synthesis of nylon 6-clay hybrid by montmorillonite intercalated with ε-caprolactam. *J. Polym. Sci., Polym. Chem. Ed.* **31**, 983–6

Kojima Y, Usuki A, Kawasumi M, Okada A, Kurauchi T, Kamigaito O 1993c One-pot synthesis of nylon 6–clay hybrid. *J. Polym. Sci., Polym. Chem. Ed.* **31**, 1755–8

Kojima Y, Usuki A, Kawasumi M, Okada A, Kurauchi T, Kamigaito O 1993d Sorption of water in nylon 6–clay hybrid. *J. Appl. Polym. Sci.* **49**, 1259–64

Krishnamoorti R, Vaia R A, Giannelis E P 1996 Structure and dynamics of polymer-layered silicate nanocomposites. *Chem. Mater.* **8**, 1728–34

Lee D C, Jang L W 1996 Preparation and characterization of PMMA–clay hybrid composite by emulsion polymerization. *J. Appl. Polym. Sci.* **61**, 1117–22

Maaref S, Frisch H L, Pu Z, Mark J E, Rajan G S, Beaucage G 1999 Composites with zeolite 13X or Vycor glass. *J. Macromol. Sci., Pure Appl. Chem.* **A36**, 1895

Messersmith P B, Giannelis E P 1994 Synthesis and characterization of layered silicate–epoxy nanocomposites. *Chem. Mater.* **6**, 1719–25

Okada A, Usuki A 1995 The chemistry of polymer–clay hybrids. *Mater. Sci. Eng.* **C3**, 109–15

Pu Z, Mark J E, Beaucage G, Maaref S, Frisch H L 1996 Small angle X-ray scattering investigations of zeolite-penetrated poly(ethyl acrylate) composites. *J. Polym. Sci., Polym. Phys. Ed.* **34**, 2657–60

Usuki A, Koiwai A, Kojima Y, Kawasumi M, Okada A, Kurauchi T, Kamigaito O 1995 Interaction of nylon 6–clay surface and mechanical properties of nylon 6-clay hybrid. *J. Appl. Polym. Sci.* **55**, 119–23

Vaia R A, Ishii H, Giannelis E P 1993 Synthesis and properties of two-dimensional nanostructures by direct intercalation of polymer melts in layered silicates. *Chem. Mater.* **5**, 1694–6

Vaia R A, Sauer B B, Tse O K, Giannelis E P 1997 Relaxations of confined chains in polymer nanocomposites: glass transition properties of poly(ethylene oxide) intercalated in montmorillonite. *J. Polym. Sci., Polym. Phys. Ed.* **35**, 59–67

Wen J, Mark J E 1994 Mechanical properties and structural characterization of poly(dimethylsiloxane) elastomers reinforced with zeolite fillers. *J. Mater. Sci.* **29**, 499–503

Wu C, Bein T 1994a Polyaniline wires in oxidant-containing mesoporous channel hosts. *Chem. Mater.* **6**, 1109–12
Wu C, Bein T 1994b Conducting polyaniline filaments in a mesoporous channel host. *Science* **264**, 1757–9
Yano K, Usuki A, Okada A, Kurauchi T, Kamigaito O 1993 Synthesis and properties of polyimide–clay hybrid. *J. Polym. Sci., Polym. Chem. Ed.* **31**, 2493–8

H. L. Frisch and L. Huang
State University of New York at Albany, New York, USA

Porous Ceramic Processing

There are numerous applications in which it is necessary or useful to introduce porosity into a ceramic material. For example, in filtration applications, materials need to be permeable and the processing emphasis is often focused on controlling the size of the pore channels. Applications include the removal of particles or large molecules from liquids (for concentration or purification) and gas separation. Similarly, many sensors, catalytic systems, and gas burners require materials to be permeable as well as providing other functions. Porosity can also be used as a second phase to design materials with tailored properties, and in this case it is necessary to control both porosity volume fraction and its size. For example, porosity can be used to decrease thermal conductivity, which is important in refractory and thermal insulation materials. The presence of porosity can decrease the dielectric constant of a ceramic, which is a property sought in substrates used for microelectronics. The void space in porous ceramics can also be used as a storage medium for the later release of the second phase. For example, this idea has been exploited for medication delivery devices or as diffusion sources in the semiconductor industry. Porous ceramics are also being used as bone implants and the porous structure has been shown to promote bone growth. The decrease of density with the addition of porosity can be used in some structures to produce materials of a given stiffness or strength with minimum weight. The sandwich panel structure, in which a porous (core) material is bonded to strong, stiff faceplates, is an example of this design methodology. The primary function of porous ceramics is, however, not often structural and the presence of porosity usually degrades strength. As many applications require a high degree of mechanical reliability, there is often a need to ensure a compromise between strength and the other properties required for the given application. Porous ceramics are often attractive in applications that involve high temperature or corrosive environments.

The processing techniques required for porous ceramics depend critically on the volume fraction and size of the required porosity. For example, ceramics with porosity levels greater than 70 vol.% can be produced with a (hollow) cellular structure. The processing of cellular ceramics often involves a foaming process and the pore size is typically greater than 100 μm. High-porosity ceramics can also be produced by the bonding or sintering of ceramic fibers. These are termed fibrous ceramics and such materials are commonly used as high-temperature thermal insulation. For example, the thermal protection system of the NASA space shuttle is based on fibrous tiles that use bonded silica or aluminosilicate fibers. An example of the microstructure of this material is given in Fig. 1. Another approach for processing high-porosity materials is to use sol–gel technology and the pore size in these materials is typically less than 1 μm. To produce ceramics with lower levels of porosity, conventional ceramic powder processing is often used but modified to ensure control of the pore volume and size. The main aspects of these various established processing approaches are discussed below. Porous ceramics can also be produced by the oxidation of metals, gas dissolution during solidification, or removal of a (continuous) second phase, e.g., by leaching. An example of the third approach is the production of “thirsty” glass (Vycor), in which one phase is leached from a phase-separated glass. This approach has also been applied to glass ceramics. It is anticipated that porous materials will also be produced by solid freeform processing or by the deposition/etching techniques that are employed in the microelectronics industry.

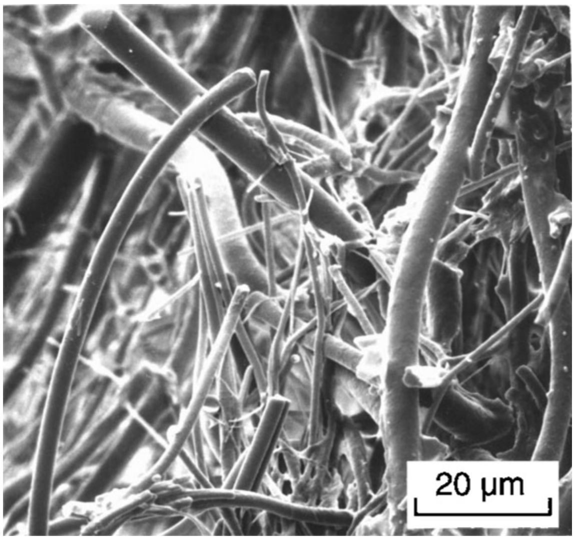


Figure 1
Fibrous structure of the high-porosity tiles used in the thermal protection system of the space shuttles.

1. Cellular Ceramics

Cellular materials are generally classified into two broad groups, honeycombs and foams. In honeycombs, the cells form a two-dimensional array whereas foams are comprised of a three-dimensional array of hollow polygons. Ceramic honeycombs have found a substantial market as catalytic converters, heat exchangers, and combustors. Foams are usually divided into two further categories depending on whether or not the individual cells possess solid faces. If the solid is contained only in cell edges, the material is termed open cell (or reticulated) and the material is permeable. These materials are often used as filters for molten metals or as kiln furniture. An example of an open-cell ceramic is shown in Fig. 2. If the cell faces are present, the foam is termed closed cell and the individual cells are isolated from each other. Closed-cell glass foams are commercially available as thermal insulation or as abrasives.

Ceramic honeycombs are usually formed by extrusion and this structure then undergoes a conventional binder burn-out and sintering process. The die used in the extrusion process can be designed to produce a wide variety of cell geometric structures, e.g., square, triangular, hexagonal, etc. The extrusion process requires moldable raw materials, the rheological behavior of which can be adjusted by water-soluble additives. In this process, a balance is needed in this rheology to ensure stability of the cellular structure without excessive die wear.

A common processing approach for open-cell ceramics is to coat open-cell polymer foams, usually with a ceramic powder dispersed in a liquid (slurry). This approach approximately replicates the polymer and gives a rather distinct structure in that the cell struts are hollow after the process is complete. The central hole represents the original shape and position of the polymer, which is removed by pyrolysis during the firing process. Various open-cell polymers have been used as substrates for the coating process including polyurethane, poly(vinyl chloride), polystyrene, and latex. The composition of the ceramic slurry used to coat the polymer depends on the final application but has included a wide variety of oxide and nonoxide ceramics. The ceramic slurry often contains various additives to act as binders, rheological control agents, or setting compounds to facilitate the coating process. It is critical that the slurry produces a uniform coating, especially as the polymer foam can often contain sharp corners at the strut edges. The presence of longitudinal strut cracks in Fig. 2 is probably a result of poor coating at the strut corners. In the processing, the polymer "sponge" is immersed in the ceramic slurry. The excess slurry is then removed by mechanical compression or centrifugation. The ceramic foam is dried and fired in a typical manner, except for a pyrolysis step in which the organic additives and polymer foam are removed. During heating, the expansion and gas evolution behavior of the polymer can lead to significant stresses that can damage the ceramic coating if not carefully controlled. For these materials, the pore size in the final product is determined primarily by the cell size of the polymer foam.

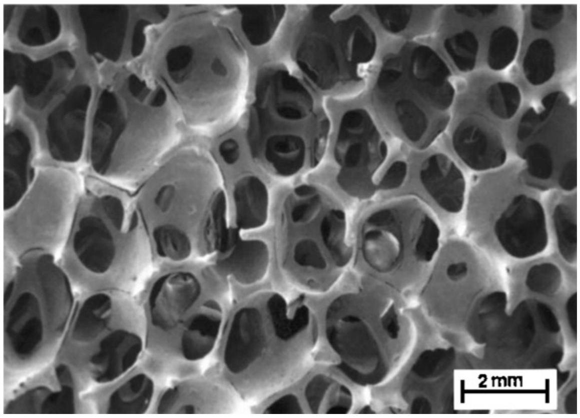


Figure 2
Structure of an open-cell ceramic produced by the sponge replication technique. This material shows the presence of longitudinal strut cracks that probably result from nonuniform coating.

An alternative approach for producing open-cell ceramics is to use a physical deposition process on an open-cell substrate. For example, open-cell vitreous carbon foams can be produced by pyrolysis of thermosetting polymer foams. The carbon foam can then be coated with a dense ceramic material, e.g., by chemical vapor deposition. This process has been used to produce a variety of oxide, boride, nitride, and carbide foams. Replication has also been performed on biological materials. For example, a lost wax process has been used to replicate coral and sols have been absorbed into wood to produce a porous structure after the firing process.

In contrast to the above replication methods, there are direct techniques to foam a liquid in which a ceramic powder or a precursor to a ceramic is dispersed. The liquid also contains a constituent which, when treated appropriately, will lead to gas evolution causing bubbles to form in the liquid. The gas evolution process is usually initiated by heat or chemical reaction and a wide variety of processes have been devised. For example, organic compounds can be mixed to produce carbon dioxide or hydrogen in an ambient temperature reaction. Other additives to the liquid are used to stabilize and harden the cellular structure once it is produced. The hardening may occur by a chemical reaction or gelation and is used to "set" the foamed structure. The foam is dried and sintered by conventional means. The final structure of a foamed material often exhibits some morphological anisotropy associated with the "rise" direction of the foam. Glass foams are produced by a high-temperature version of this process in which additives such as hydroxides, carbonates, and sulfates are mixed with glass powder. On heating, the glass becomes fluid and the decomposition of the salts leads to bubble formation. The direct foaming approaches give rise to a closed-cell structure, but if sintering is used the cell faces often perforate and produce an open-cell ceramic. As an alternative to direct foaming, the sintering or bonding of hollow spheres has been used to produce closed-cell ceramics.

2. *Fibrous Ceramics*

Fibers in a blanket or felt form have been used extensively as thermal insulation but fibers can also be "rigidized" by bonding processes to produce fibrous ceramic monoliths (tiles or boards). These materials are used for a wide variety of thermal insulation applications and are attractive as a result of their low thermal mass, which allows fuel savings. These bodies are usually classified in terms of the refractoriness of the fibers. They may range from low- to high-purity silica glasses or from low- to high-purity alumina fibers, depending on the alumina/silica ratio in the latter fibers. The low-alumina fibers are usually produced by a melt blowing process. One of the processing concerns in blown fibers is the amount of "shot" (solidified particles that escape the fiber-shaping process) that is present in the final product. High-alumina fibers (ICI's Saffil alumina or 3M Corporation's Nextel mullite) are produced by a sol–gel route and can be used at temperatures up to 1600 °C.

Fibrous ceramics are often formed by a uniaxial pressing operation and this can lead to significant anisotropy in the physical properties. In this process, conditioned fibers are mixed with a liquid that contains the appropriate binders and dispersants. This slurry is then mechanically pressed or molded into the required shape and gelation is used to set the structure. The green body is dried and fired by conventional means, although the low thermal conductivity of these materials can lead to longer processing times being required to reach thermal equilibrium.

3. *Porous Sol–Gel-derived Ceramics*

Sol–gel chemistry can be used to produce porous ceramics either as monoliths or as thin films. The latter are termed ceramic membranes and are used for filtration. The pore size in these sol–gel-derived materials can be controlled from approximately 1 nm to 1 μm. The sol–gel process can be accomplished using a wide variety of techniques. For example, alkoxides such as tetraethyl orthosilicate can be mixed with water to initiate the gelation. Subsequently, hydrolysis, condensation, and polymerization take place to form an alcogel, which is an oxide network that has condensed with liquid (alcohol) filling the porous network. The liquid is then removed by evaporation (xerogels) or by hypercritical drying (aerogels) to produce high-porosity silica. The pore structure can then be further modified by a heat treatment. Alumina, zirconia, and titiania ceramics can be produced by similar sol–gel processes. An alternative sol–gel process is one in which a colloidal suspension of hydoxide or hydrated oxide particles is peptized. Gel formation is induced by increasing particle concentration or by manipulating the zeta potential. The process is controlled by pH, the nature/concentration of the electrolyte, and particle size.

Zeolites are materials in which porosity is available at the molecular level and sol–gel technology has extended this behavior to larger pore sizes. An interesting development of these ideas is the production of porous silica-based material, in which the pores are ordered at the nanoscale. For example, Kresge *et al.* (1992) produced silica-based materials in which the pores are in hexagonal arrays and the pore sizes are in the range 1.6–10 nm. This has proved to be a very active research area and scientists have worked to extend the pore size range and choice of materials. These materials use organic "templates," such as surfactant arrays or emulsion droplets that allow the porous materials to be self-assembled. Latex spheres and block copolymers can also be used to introduce submicrometer pores into sol–gel-derived materials, and it has been demonstrated that three-dimensional ordered structures can be produced by such means.

4. *Partially Densified Ceramics*

Many ceramics are produced by the (high-temperature) firing of powder compacts. The level of porosity in these materials can be controlled by either firing at a temperature at which densification is incomplete (see Fig. 3) or by controlling the particle size distribution. This latter approach is used extensively in the refractories industry. The microstructure of refractories usually consists of grain and bond phases. The large refractory grains are often derived from the raw materials with a minimum of preprocessing. The bonding phase is usually fine-grained, such as the glassy or silicate bonding derived from clays, carbon bonding derived from resin or pitch, or calcium aluminate cements. During the powder processing, the fractions of the coarse, intermediate, and fine-sized particles are carefully controlled as this defines not only the green density but also the amount of shrinkage that occurs during firing. Unlike other engineering ceramics, refractories may only have limited shrinkage during firing. The firing process is, however, important in controlling the bonding process,

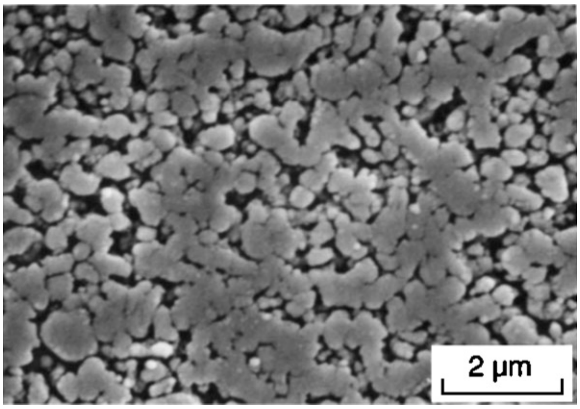


Figure 3
Porous alumina ceramic produced by the incomplete densification of an ultrafine powder (courtesy of Dale Hardy).

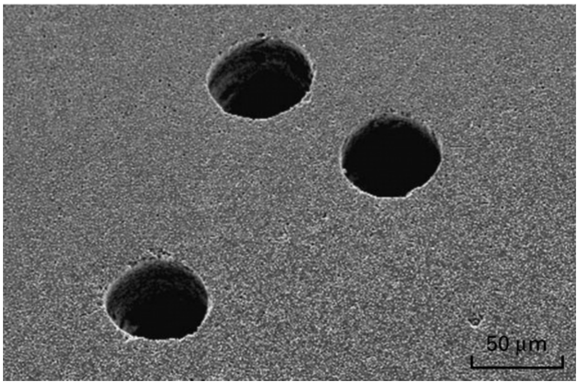


Figure 4
Example of a porous ceramic in which the pores are introduced by the pyrolysis of polymeric fibers. The fibers are aligned to some degree by a tape-casting process (courtesy of Jordan Priest).

which leads to strength development. As refractories can contain several phases, there is the possibility of residual stress development from thermal expansion mismatch and the molar volume changes that arise from the various chemical reactions. Clearly, these stresses must be controlled so that any microcracking does not have an impact on the desired properties in the final product.

Porosity can also be introduced into powder-processed ceramics by adding a fugitive phase to the powder mix. For example, inexpensive raw materials, such as sawdust or starch, are employed. These organic materials are "burnt off" by the firing process leaving the additional void space. Alternatively salts, such as carbonates or hydoxides, which decompose by gas evolution, can be employed. If porosity of well-defined shape is required, polymeric fibers or particles can be used. For example, Fig. 4 shows a porous ceramic in which the pores are produced by the pyrolysis of polymeric fibers. Some ceramic materials are available in the form of hollow spheres, and incorporation of these materials into a powder mix can produce well-defined porosity. In the cement industry, hollow flyash spheres (cenospheres) are sometimes used to produce a lightweight cement or concrete.

See also: Glasses: Sol–Gel Methods[#]; Porous Ceramics including Fibrous Insulation, Structure and Properties of[#]

Bibliography

Brezny R, Green D J 1994 Mechanical behavior of cellular ceramics In: *Materials Science and Technology—A Comprehensive Treatment*. VCH, Weinheim, Germany, Vol. 11 Chap. 9

Brinker C J, Scherer G W 1990 *Sol–Gel Science: The Physics and Chemistry of Sol–Gel Processing*. Academic Press, New York

Burggraaf A J, Cot L (eds.) 1996 *Fundamentals of Inorganic Membrane Science and Technology*. Elsevier, New York

Chan K K, Brownstein A M 1991 Ceramic membranes—growth prospects and opportunities. *Bull. Am. Ceram. Soc.* **70**, 703–7

Gibson L J, Ashby M F 1997 *Cellular Solids: Structure and Properties*. Cambridge University Press, Cambridge

Hirschfeld D A, Li T K, Liu D M 1996 Processing of porous oxide ceramics. *Key Eng. Mater.* **115**, 65–79

Holland B T, Blanford C F, Stein A 1998 Synthesis of macroporous minerals with highly ordered three-dimensional arrays of spheroidal voids. *Science* **281**, 538–40

Kresge C T, Leonowicz M E, Roth W J, Vartuli J C, Beck J S 1992 Ordered mesoporous molecular sieves synthesized by a liquid-crystal template mechanism. *Nature* **359**, 710–2

Nettleship I 1996 Applications of porous ceramics. In: Mostaghaci H (ed.) *Applications of Advanced Materials in a High-Tech Society: Part 1* Trans Tech, Zurich, Switzerland, pp. 305–24

Pierre A C 1998 *Introduction to Sol–Gel Processing*. Kluwer, Boston

Saggio-Woyansky J, Scott C E, Minnear W P 1992 Processing of porous ceramics. *Bull. Am. Ceram. Soc.* **71**, 1674–82

Schaefer D W 1994 Engineered porous materials. *MRS Bull.* **19**, 14–7

Sepulveda P 1997 Gelcasting foams for porous ceramics. *Bull. Am. Ceram. Soc.* **76**, 61–5

Shermann A J, Tuffias R H, Kaplan R B 1991 Refractory ceramic foams—a novel new high-temperature structure. *Bull. Am. Ceram. Soc.* **70**, 1025–9

D. J. Green
Pennsylvania State University, University Park, Pennsylvania, USA

Porous Silica Xerogel

Scaling of integrated circuit (IC) feature sizes has brought tremendous performance improvement and miniaturization. The primary contribution to IC performance improvement has been from the device gate level. However, metal interconnects become the limiting factors for continued IC performance improvement when the feature size scales down to the deep-submicrometer region and below. The gain in device speed at the gate level is offset by propagation delays at metal interconnects owing to the increased RC time constant. The RC time delay can be reduced by the incorporation of low dielectric constant (k) materials and/or high conductivity metals. The use of low-k dielectric materials also lowers power consumption and reduces crosstalk (Jeng *et al.* 1994).

There are numerous low-k dielectric candidates available with dielectric constants roughly in the range 2.4–4.0 (Lee and Ho 1997). As predicted by the Semiconductor Industry Association (1998), continued scaling of devices will require ultralow-k materials with dielectric constants as low as 1.5. One of the ways to obtain a lower dielectric constant is fluorination of either inorganic or organic dielectric materials. However, the lowest dielectric constant available for fluorinated dense materials is $k \sim 1.9$ (Teflon) and none of the approaches using dense materials is expected to achieve k values lower than this. Air practically has the lowest dielectric constant possible, $k \sim 1$. Incorporation of air into dense materials to make them porous is another attractive method to obtain ultralow-k materials. The ultralow dielectric constant results from the incorporation of pores. For a porous material, the dielectric constant is a combination of air and that of the dense phase, and it has the potential to drive the dielectric constant significantly below 2. To meet performance goals, future technology nodes will require materials with progressively lower dielectric constant. A change in dielectric material for each technology node increases process and equipment complexity and development costs. It is desirable to have one kind of material that can meet the requirements for multiple future technology nodes. Porous dielectric materials offer the extendibility to multiple technology nodes because they have a tunable dielectric constant. The dielectric constant variation with respect to porosity (volume fraction of pores) of porous silica xerogel is shown in Fig. 1. The figure includes predictions from the parallel model as well as bulk sample measurements (Hrubesh *et al.* 1993). Clearly, a dielectric constant of interest can be achieved by tuning the porosity to a proper level. Besides the tunable k value, porous silica xerogel also has the advantage of using SiO_2 as the base material. It offers some of the advantages SiO_2 has in process integration such as high thermal stability, etc.

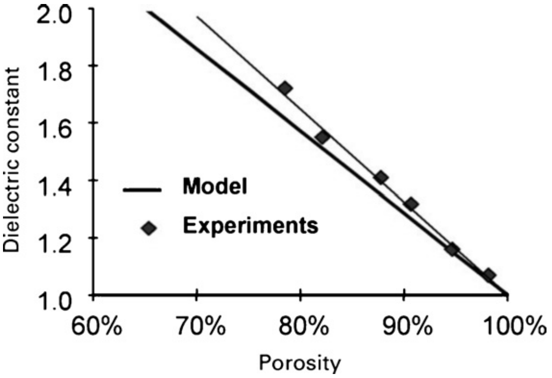


Figure 1
Dielectric constant of porous silica xerogel as a function of porosity based on a parallel model and experiments on bulk samples.

1. Porous Silica Xerogel Processing

Porous silica xerogel is produced using sol–gel processes through hydrolysis and polycondensation of an alkoxide, such as tetraethoxysilane (TEOS), followed by drying of the gel to form a porous silica network (Iler 1979, Brinker and Scherer 1990). Porous silica xerogel films for semiconductor applications are deposited using spin coating. Figure 2 depicts an overall process block diagram (Jin *et al.* 1996, Ramos *et al.* 1996).

In precursor preparation, the hydrolysis and condensation reactions are controlled to obtain solutions suitable for spin coating. The precursor is then dispensed onto a substrate and spin-coated to produce a uniform film. If the precursor is dispensed and coated in an open environment typical of most spin-on-dielectric (SOD) processes, polycondensation, gelation, and solvent evaporation stages can overlap during spin coating. Ethanol is usually used as the common solvent for water and TEOS in the sol–gel process. Severe solvent evaporation can occur during spin coating owing to the high vapor pressure of ethanol. It causes premature drying and shrinkage of the gel. To control the solvent evaporation, a sealed chamber saturated with ethanol vapor is necessary for precursor dispense and coating, making the process neither safe nor practical for manufacturing. To solve this problem, alternative low vapor pressure solvents are used to reduce the solvent evaporation during coating. The precursor now can be dispensed and coated without ambient control on regular commercial SOD process equipment.

After spin coating, the wafer is aged in a controlled environment with NH_4OH vapor. Aging of the silica gel is necessary to increase its mechanical strength and prevent unnecessary shrinkage in a subsequent drying step. Again, solvent evaporation from the film

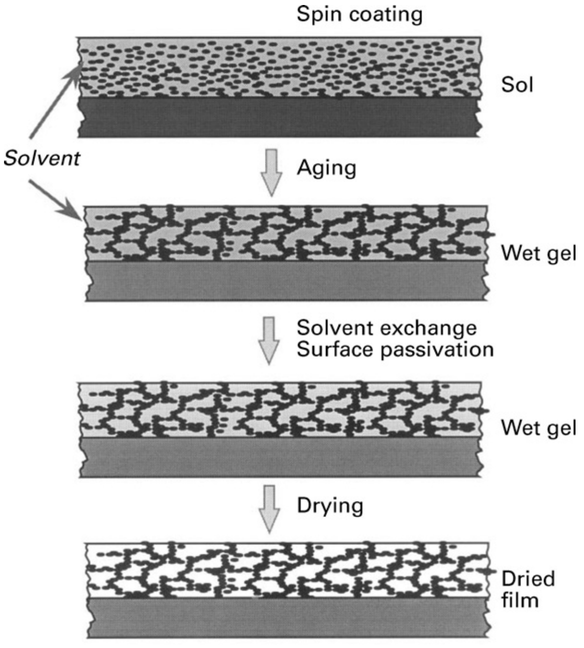


Figure 2
Porous silica xerogel film deposition steps using sol–gel processes by spin coating.

is minimized during aging. After aging, the original solvent is exchanged with a low surface tension solvent to reduce the capillary pressure in micropores during drying. Because of the high porosity and small pore sizes, silica xerogel has a large surface area, typically in the range 500–1000 $m^2\,g^{-1}$. Hydroxyl groups usually terminate the surface silicon atoms. The average coverage is around 4.9 $OH\,nm^{-2}$ and is largely independent of process conditions. These surface hydroxyl groups are the primary sites for moisture adsorption, which increases the dielectric constant and causes integration problems. The surface silanol groups also form Si–O–Si bonds through condensation reactions during drying, resulting permanent shrinkage of the film. To reduce moisture adsorption and film shrinkage during drying, surface hydroxyl groups in xerogel films are replaced with noncondensing organic groups such as methyl in a surface passivation step. Since no permanent Si–O–Si bonds are formed, the film can "spring-back" after drying. Surface passivation is typically conducted during the solvent exchange step. The xerogel film is then dried on a series of hot plates and sometimes thermally annealed in a furnace.

The dielectric constant of porous silica xerogel films can be tuned by adjusting the solid content in the precursor. The aging and surface passivation steps are critical in controlling xerogel film properties. The dielectric constant can easily vary by more than 20% for the same chemical precursor under different aging and/or surface passivation process conditions.

2. *Porous Silica Xerogel Film Characterization and Properties*

Porous silica xerogel films have been partially characterized for semiconductor applications (Jin *et al.* 1996). The dielectric constant measured using the metal–insulation–semiconductor capacitor method is $k \sim 1.8$–2.5 depending on porosity and process conditions. Porous silica xerogel films display good dielectric strength with a breakdown field in excess of 2 $MVcm^{-1}$. Because the pore sizes of porous silica xerogel films are typically less than 10 nm and are much smaller than the mean free path (~60 nm) of ambient air, the film dielectric strength is much greater than that of ambient air. Porosity of the films is measured using spectroscopic ellipsometry and Rutherford backscattering spectroscopy (RBS). Ellipsometry data fit very well with a uniform SiO_2 and air mixture model. RBS measures the area density of the film, and the porosity is calculated based on the film thickness from ellipsometry measurements and the SiO_2 density of 2.22 $g\,cm^{-3}$. Porosities obtained from RBS and ellipsometry agree with each other. Porosity can also be measured using specular x-ray reflectivity (SXR). In SXR, electron density is directly determined from the critical angle where an x-ray beam is no longer totally reflected from the film surface. If chemical composition of the material is known, the mass density can be calculated. The porosity can then be derived in the same way as for RBS. Adhesion of porous silica xerogel films to silicon and plasma enhanced TEOS-based CVD oxide (PETEOS) substrates are analyzed using Scotch tape and stud pull tests. In both cases, no delamination occurs between the porous silica xerogel films and the substrates. Porous silica xerogel films exhibit good adhesion to both silicon and PETEOS substrates. However, cohesive failures can occur if the xerogel films themselves are not strong enough. As expected for highly porous materials, porous silica xerogel films display very low stress in the temperature range 20–500 °C. FTIR spectra of the films show an absence of moisture peaks after the surface passivation step. The surface contact angle can reach nearly 90° in water drop contact angle measurements, confirming the hydrophobicity of the xerogel films.

One of the major drawbacks of porous materials is their mechanical strength. Just as the dielectric constant of a porous material is between that of its dense phase and that of air, mechanical properties such as Young's modulus also fall between that of its dense phase and of air. Young's modulus of a porous material depends strongly on its density and usually exhibits a power law relationship, $K \sim (\rho/\rho_0)^m$, where

K is Young's modulus, ρ and ρ_0 are the porous material density and a reference density, respectively, and $m = 3\text{–}4$. Changing the density, and therefore the porosity, will have a significant impact on the mechanical properties of a porous material. In addition, process conditions also play an important role in determining the mechanical properties. An order of magnitude change in modulus can be obtained by varying process conditions.

3. *Integration of Porous Silica Xerogel Films*

The feasibility of integrating porous silica xerogel films into IC test structures has been studied for both aluminum and copper metallization processes (Zielinski *et al.* 1997, Jin *et al.* 1996). In conventional non-damascene aluminum metallization processes, gapfill capability is an important factor for intermetal dielectrics. Because low vapor pressure solvent is used in the xerogel chemical precursor, the film remains wet and flowable after spin coating giving sufficient time for it to fill complex gap structures. However, shrinkage during aging and drying may cause the porous silica xerogel film to crack and pull back from narrow gaps between metal lines. With proper process control and without resorting to an elaborate process setup and/or supercritical drying, shrinkage can be minimized such that narrow gaps with high aspect ratios can be filled.

A dense capping layer on top of porous silica xerogel films is needed for chemical mechanical polish (CMP) and etch processes. Various CVD dielectric films are deposited on top of porous silica xerogel films without deeply penetrating into the pores and changing the porosity of the underlining films. However, because xerogel pore surfaces are passivated with organic groups such as methyl, they are susceptible to oxidizing plasma damage. The removal of the organic passivation groups such as methyl will result in high silanol content and increase moisture adsorption. Optimized capping layer deposition, etching, photoresist ashing, and cleaning processes are needed to minimize the damage. FTIR is usually used as a quick way to monitor damages to low-*k* films. One can also measure other parameters such as dielectric constant, moisture adsorption, etc. A reducing atmosphere instead of the conventional O_2 plasma may be used in photoresist ashing. C_2F_6-type chemistry has been tested in a plasma reactor to etch porous silica xerogel films. Porous silica xerogel films etch much faster than dense CVD oxides under the same conditions although the etch rate is controllable. Vias and trenches etched through these porous films have reasonably vertical profiles.

Concerns over their mechanical strength require compatibility studies of porous silica xerogel films with processes that generate high mechanical stress such as CMP and CVD tungsten plug deposition. To integrate xerogel films with the CMP process, they are capped with a dense CVD oxide film. To prevent plasma damage and ensure mechanical integrity of the film stack, a buffer layer between the xerogel film and the CVD cap oxide may be incorporated. The CVD oxide film is successfully polished, indicating the feasibility of the CMP process for such a film stack. The xerogel film maintains its porosity and thickness after CMP. CVD tungsten is also deposited on such a film stack without delamination.

List *et al.* (1997) have successfully integrated porous silica xerogel films into double-level metal (DLM) test structures using aluminum metallization processes. High-density plasma (HDP) CVD oxide and spin-on hydrogen silsesquioxane (HSQ) are used as controls. Both of these materials are used in semiconductor manufacturing. Metal lines with 0.3 μm width/spacing are patterned and etched. Part of the wafers receives an HDP oxide deposition. For the rest of the wafers, a 50 nm CVD oxide layer with about 50% sidewall coverage is deposited as a liner to cover the aluminum metal lines before the coating of xerogel and HSQ. After deposition, xerogel and HSQ films are capped with a CVD oxide film. A buffer layer is used between the xerogel and the CVD oxide cap. All wafers are then planarized using CMP. Splits are run to fill vias with aluminum and tungsten plugs, respectively, after via pattern and etch. All wafers receive a final 450 °C sinter.

Electrical measurements and analysis are performed to determine total capacitance, leakage current, via resistance, and electromigration lifetimes in structures with xerogel, HSQ, and HDP oxide dielectric materials. 0.3 μm line/space interdigited comb structures are used for the total capacitance distribution measurements. The xerogel wafers demonstrate a total capacitance reduction of 35% compared to HDP oxide and 14% reduction compared to HSQ. Raphael modeling of these structures gives a fitted xerogel dielectric constant of 2.0. None of the materials shows any frequency dependence in capacitance between 1 kHz and 1 MHz. All the structures have very low leakage currents and via resistance. As shown in Fig. 3, reliability measurements on bow tie metal leads indicate that aluminum/xerogel structures have similar electromigration lifetime as those of aluminum/HSQ structures.

Porous silica xerogel has also been integrated into copper damascene structures (Zielinski *et al.* 1997). In the copper/xerogel damascene process flow, xerogel films are coated on substrates followed by the deposition of a buffer layer and a CVD oxide cap. The buffer layer serves the purpose of improving film stack integrity. 0.3 μm trenches are patterned and etched. After deposition of PVD barrier and copper seed layers, the trenches are filled with electroplated copper. Copper lines are formed by CMP removal of copper on the field with a minimal amount of oxide loss and copper dishing in wide

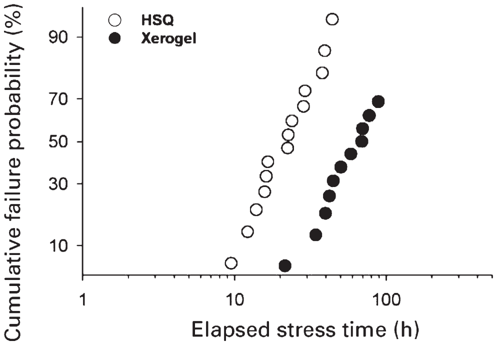


Figure 3
Distribution of electromigration failures for porous silica xerogel and HSQ in aluminum gapfill structures measured on bow tie metal leads.

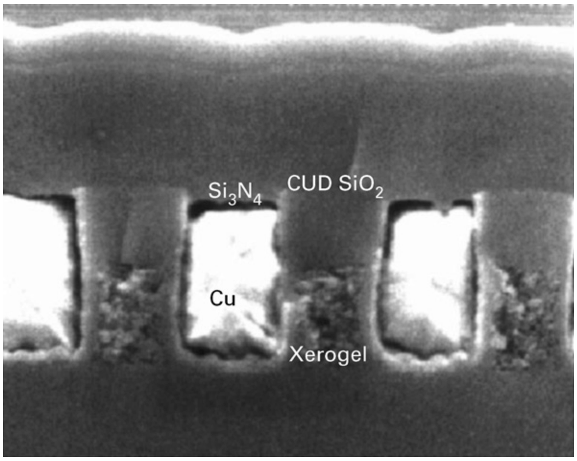


Figure 4
Cross-sectional SEM image of copper lines embedded in porous silica xerogel trenches.

lines. Copper is then capped with a silicon nitride/oxide/silicon nitride stack and bond pads are opened followed by electrical measurements. Figure 4 shows a cross-sectional scanning electron microscopy (SEM) image of copper lines embedded in xerogel trenches. In comparison with a standard aluminum/oxide structure, the copper/xerogel structure shows advantages both in measured total capacitance and metal line resistance.

4. Conclusions

It is generally believed that to drive the dielectric constant below 2 while satisfying integration requirements, incorporation of porosity into dielectric materials is necessary. Porous silica xerogel is the first porous material being proposed as an ultralow-*k* dielectric and is the most studied. Numerous porous dielectrics based on different materials and obtained through various approaches are becoming available. One of these approaches is to deposit a composite of thermally labile organic polymers with thermally stable materials. Upon thermal treatment under various process conditions, the labile components are removed leaving pores in the stable material. Porous materials that have been produced this way include surfactant templated porous silica, block polymers, and organic/inorganic composite materials (Bruinsma *et al.* 1996, Miller *et al.* 1997). Although there are many material, processing, metrology, and integration issues to be resolved, preliminary studies have shown the feasibility of incorporating porous materials into ULSI test structures.

Bibliography

Brinker C J, Scherer G W 1990 *Sol–Gel Science*. Academic Press, San Diego, CA

Bruinsma P J, Hess N J, Bontha J R, Liu J, Baskaran S 1996 Low *k* mesoporous silica films through templated-based processing. In: Lagendijk A, Treichel H, Uram K, Jones A (eds.) *Low-Dielectric Constant Materials II*. MRS, Pittsburgh, PA, pp. 105–10

Hrubesh L W, Keene L E, Latorre V R 1993 Dielectric properties of aerogels. *J. Mater. Res.* **8**, 1736–41

Iler R K 1979 *The Chemistry of Silica*. Wiley, New York

Jeng S-P, Chang M, Kroger T, McAnally P, Havemann R H 1994 A planarized multilevel interconnect scheme with embedded low-dielectric-constant polymers for sub-quarter-micron applications. *1994 VLSI Tech. Dig.*, 73–4

Jin C, List R S, Lee W, Lee C, Luttmer J D, Havemann R H 1996 Porous xerogel films as ultra-low permittivity dielectrics for ULSI interconnect applications. In: Havemann R H (ed.) *Advanced Metallization for ULSI Applications*. MRS, Pittsburgh, PA, pp. 463–9

Lee W W, Ho P S 1997 Low-dielectric-constant materials. *MRS Bull.* **22**, 19–24

Miller R D, Hedrick J L, Yoon D Y, Cook R F, Hummel J P 1997 Phase-separated inorganic–organic hybrids for microelectronic applications. *MRS Bull.* **22**, 44–8

Ramos T, Roderick K, Maskara A, Smith D M 1996 Nanoporous silica for low-*k* dielectrics. In: Lagendijk A, Treichel H, Uram K, Jones A (eds.) *Low-Dielectric Constant Materials II*. MRS, Pittsburgh, PA, pp. 91–8

Semiconductor Industry Association 1998 *International Technology Roadmap for Semiconductors 1998 Update*. SEMATECH, Austin, TX

Zielinski E M, Russell S W, List R S, Wilson A M, Jin C, Newton K J, Lu J P, Hurd T, Hsu W Y, Cordasco V, Gopikanth M, Korthuis V, Lee W, Cerny G, Russell N M, Smith P B, O'Brien S, Havemann R H 1997 Damascene integration of copper and ultra-low-*k* xerogel for high performance interconnects. *1997 IEDM Tech. Dig.*, 936–8

C. Jin
Texas Instruments, Dallas, Texas, USA

Portland Cements

Cement is the binder used to make concrete mortars, stuccos and grouts for all types of building and construction. Although cement is a generic term and can be applied to many inorganic and organic materials, by far the most widely used and most versatile cement is Portland cement. This article is primarily concerned with Portland cement, but a brief description of other inorganic cements is also given. Unlike gypsum plaster, Portland cement is a hydraulic cement; i.e., it sets and hardens under water and hence can be used safely in all structures in contact with water. This article describes the manufacture and chemical composition of Portland cements used by the construction industry. It also summarizes the chemistry and microstructure of hydrated cement paste which is the binder in concrete and masonary construction.

The ancestry of Portland cement can be traced back to Greek and Roman times, but the year 1824 is generally considered to mark the origin of modern Portland cements when the name "Portland" was used in a patent by Joseph Aspdin to emphasize the similarity of concrete made with his cement to Portland limestone, which was a favored building material at the time. By the 1860s the Portland-cement industry was flourishing in the UK and Europe, but the first patent in the US was not taken out until 1870 by David Saylor. Today about 90 Mt are produced annually in the US.

1. Manufacture

Portland cement is commonly made from limestone and clay, two of the most widespread and abundant materials. The principal constituents are thus lime (CaO) and silica (SiO_2) which, when subjected to high temperatures (~1450 °C), form the hydraulic calcium silicates that are the active compounds in cement. Other calcareous materials (such as muds, marls, and shell deposits) or argillaceous materials (such as slates, shales, and coal ashes) are used in some cases.

The overall manufacturing process is shown schematically in Fig. 1. The various steps are: (i) quarrying and crushing the raw materials, (ii) proportioning and fine grinding of the crushed materials, (iii) burning in a rotary kiln at high temperature, and (iv) grinding the resulting clinker to a fine powder with the addition of a small amount of gypsum. Although wet milling and proportioning with the aid of slurries (the *wet process*) is traditional and is found in older plants in the US, modern plants all use dry grinding and proportioning (the *dry process*). Although it was made possible by improvements in grinding mills, in particle-size classification, and in dust collection, the adoption of the dry process has been motivated largely by energy efficiency.

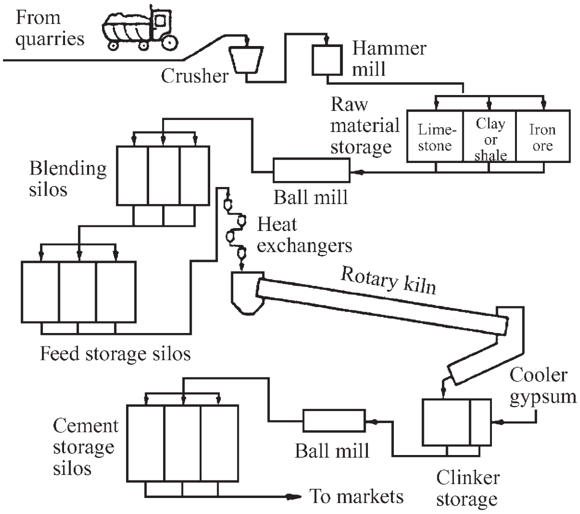


Figure 1
Schematic layout of a modern dry-process cement plant.

In the wet process, large amounts of energy are required to evaporate the water, and the long kilns, which must be used to attain large throughputs of materials (Fig. 2), have high heat losses. In addition, the dry process lends itself to the use of more efficient heat exchangers, such as the suspension preheater, in which the incoming kiln feed draws heat from the exit gases in a series of cyclones. As a result, dry-process rotary kilns are much shorter than those used in the wet process (Fig. 2), heat losses are much reduced, and throughputs are higher. Further savings can be realized by precalcining the hot feed in a flash furnace (precalciner) before it enters the kiln.

Typical energy consumptions in cement manufacture are given in Table 1. Average overall energy consumptions ($MJ\,kg^{-1}$) for countries with cement industries employing various processes are as follows; Japan, 4.5; FRG, 4.2; and US, 5.7. Theoretically a value of $1.7\,MJ\,kg^{-1}$ is calculated from the basic chemical processes involved; in practice, modern plants now have energy consumptions approaching $3.5\,MJ\,kg^{-1}$. The average figure for the US is relatively high because of a preponderance of older plants built when energy was inexpensive.

As calcination of the raw materials progresses, chemical combination takes place in the hotter zones of the rotary kiln (see Fig. 2). Formation of intermediate compounds by sintering occurs first, until at 1335 °C a quarternary liquid phase appears with a composition close to that predicted from the C-S-A-F phase diagram. ($C = CaO$, $S = SiO_2$, $A = Al_2O_3$, and $F = Fe_2O_2$ in conventional cement chemists' notation; other symbols employed include $S = SO_2$, $H = H_2O$, and $K = K_2O$.) The amount of liquid formed depends on the amount of A and F present in the raw feed and on the kiln temperature.

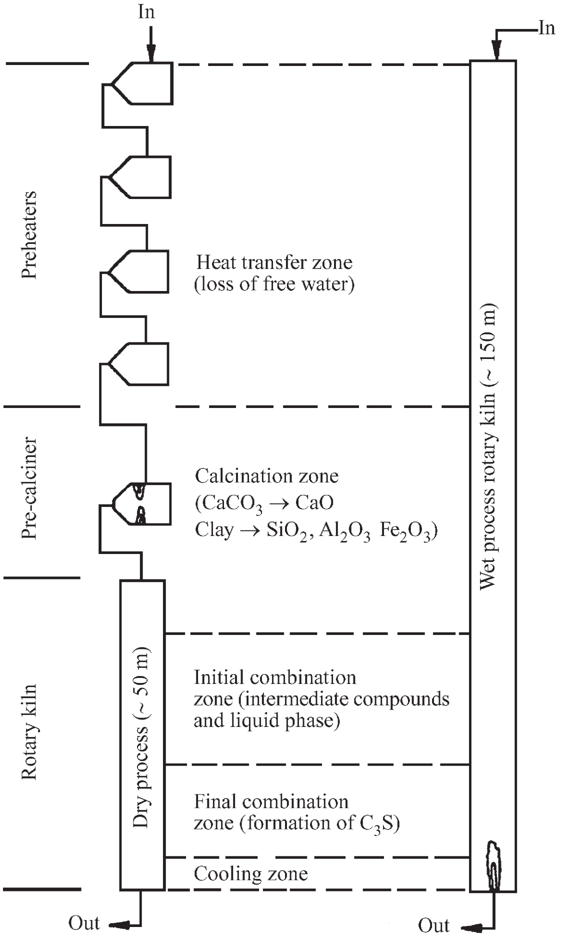


Figure 2
Comparison of kiln reactions in the wet and dry processes.

Table 1
Average energy consumption in cement manufacture ($MJ\,kg^{-1}$).

Kiln operation	Wet process	Dry process
Fuel consumption	6.5	4.7
Electrical energy	0.5	0.6

Generally, about 25% of the total mix is liquid and hence the process is neither a complete fusion or a pure solid-state sintering; it is called *clinkering*. The liquid phase allows chemical combination to go to completion in a relatively short time. Much C_2S has formed by solid-state reactions before the liquid appears but its formation is completed in the liquid. The final reaction between C_2S and lime to form C_2S takes place wholly within the liquid at 1450–1500 °C. The rate of combination depends on the rate of dissolution of lime in the clinker liquid, which is dependent on both particle size and temperature.

On leaving the hottest zone, the clinker cools to about 1100–1200 °C. Then it is air-quenched as it exits the kiln. However, cooling is usually too rapid to allow true equilibrium to be maintained. The rate of cooling also controls the crystal growth of the calcium silicates and the crystallization of liquid phase to form C_3A, C_4AF, and additional C_2S ($C_3A = 3CaO.Al_2O_3$, etc.), slow cooling results in the formation of larger and less-reactive crystals. The structure of a clinker can be studied by optical microscopy and used to predict clinker performance. After cooling, the clinker may be stored for future grinding or it may be ground immediately. Finish grinding takes place in ball mills equipped with particle-size classifiers that avoid over-grinding. At this time, small quantities of gypsum are added to control the early hydration reactions of C_3A and thus optimize cement performance.

2. *Composition*

Portland cement consists of four major clinker compounds, C_3S, C_2S, C_3A, and C_4AF together with the gypsum added during grinding. The relative portions of these compounds can be changed to optimize specific properties of the cement (discussed in more detail below). Table 2 shows typical compound compositions, which are calculated assuming (incorrectly) ideal stoichiometries and chemical equilibrium in the kiln (the Bogue calculation).

All the clinker compounds form solid solutions with small amounts of every element that is in a clinker. In addition to the four major oxides, MgO, K_2O, and Na_2O are ubiquitous impurities in most raw materials. Impure C_3S is called *alite*: it reacts more rapidly with water than does pure C_2S. C_2S exists as four major polymorphs: γ, β, α', and α. At ambient temperatures, the nonhydraulic γ polymorph is thermodynamically stable but, in cement, impurities stabilize the hydraulic β polymorph. Impure C_3A contains considerable quantities (~ 10 mol.%) of iron, which may make it less reactive than the pure compound. High levels of sodium substitution modify its crystal structure from orthorhombic to cubic with a corresponding decrease in reactivity.

C_4AF does not have an exact stoichiometry; continuous solid solution between A and F allows its composition to vary from about C_6A_2F to C_6AF_2. For this reason C_4AF is usually called the ferrite phase. Although in many cements the composition of the ferrite phase is close to C_4AF, in those that have very low C_3A contents, the composition is closer to C_6AF_2. As the iron content increases, the reactivity of the ferrite phase decreases. It is the ferrite phase that

Table 2
Average compound composition of Portland cement (wt.%).

ASTM type	Description	C_3S	C_2S	C_3A	C_4AF	$C\bar{S}_2$
I	General purpose	55	17	10	7	6
II	Moderate sulfate resistant	55	20	6	10	5
III	High early strength*	55	17	9	8	7
IV	Low heat of hydration	35	40	4	12	4
V	Sulfate resistant	55	20	4	12	4

Note: *This cement is usually ground more finely: ~550 $m^2\,kg^{-1}$ compared to 380 $m^2\,kg^{-1}$.

Table 3
Hydration of clinker compounds.

Compound	Hydration reactions	Heat of reaction	
		($kJ\,mol^{-1}$)	$J\,g^{-1}$
Alit	$2C_3S + 11H \rightarrow C_3S_2H_8 + 3CH$	−114	−500
Belite	$2C_2S + 9H \rightarrow C_3S_2H_8 + CH$	−43	−250
C_3A	$C_3A + 3C\bar{S}H_2 + 26H \rightarrow C_6A\bar{S}_3H_{32}$ $2C_3A + C_6A\bar{S}_3H_{32} + 4H \rightarrow 3C_4A\bar{S}H_{12}$	−365	−1350
Ferrite	$C_4AF + 15H \rightarrow C_4(A_3F)\bar{S}H_{12} + (A,F)H_3$	−403	−200

gives cement its color: when MgO is present in the ferrite phase, it becomes black and the cement gray.

If the MgO content is low, or if the ferrite phase contains appreciable ferrous iron (due to a reducing atmosphere in the kiln), the cement is browner. In addition, crystalline MgO (periclase) is present in most clinkers, typically in the range 2–4 wt.%. There are limits placed on the amount of MgO, because its slow hydration can cause disruptive expansions in the hardened paste. In clinkers that are produced using high sulfur coal as a fuel, alkali sulfates, such as $K\bar{S}$ or $C_2K\bar{S}_3$ sometimes occur; these compounds may affect setting behavior and can lower the 28-day strength.

Cements can be classified according to the ASTM types given in Table 2. These classifications are based mostly on performance, not composition, although in general cements of the same type have broadly similar compositions. Over 90% of all Portland cement made in the USA is type I. When higher strength is required at early ages (1–3 days) a type III can be used. This is more finely ground so that its overall hydration is faster. Type IV cement can be used to avoid thermal cracking in mass concrete, by lowering the heat of hydration and hence the internal temperature of the mass but are no longer manufactured in the US. In the presence of high concentrations of sulfates in ground waters or soils, the hydration products of C_3A may undergo further reactions that result in disruptive expansion. The resistance of concrete to sulfate attack can be improved by lowering the C_3A content in a type V cement. Type II cement has properties intermediate between those of type I and type V.

Aside from the five ASTM types of Portland cement, other special cements are available. Blended cements, containing reactive siliceous materials (pozzolans), or hydraulic slags mixed with type I, are used in most countries as alternatives to types IV and V. On hydration, they form more C–S–H and less calcium hydroxide. Expansive cements are modified Portland cements that counteract the effects of drying shrinkage by forming large amounts of ettringite to generate restrained expansion. Some special cements use ettringite to provide high early strength. Masonry cements and oil-well cements are Portland cements, modified to optimize the special requirements of these applications. Calcium aluminates (high-alumina cements) are used principally for refractory concretes. They cannot be used for structural purposes because of potential strength loss during service. Magnesium oxychloride (Sorel) cements can be used only for interior applications because they do not resist moisture.

3. *Hydration*

The setting and hardening of concrete is the result of complex chemical reactions between cement and water (hydration) in the cement paste binder. The overall reactions are summarized in Table 3. Both alite and belite form the same hydration products: an amorphous calcium silicate hydrate ($C_3S_2H_8$)

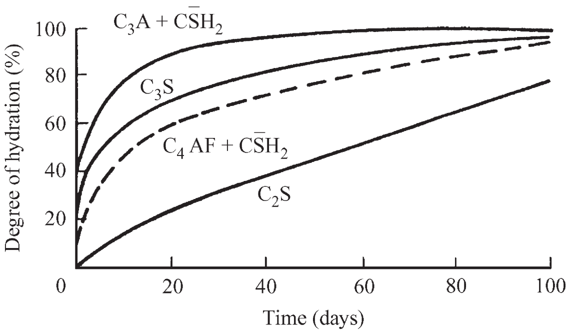


Figure 3
Rate of hydration of the clinker compounds.

and crystalline calcium hydroxide (CH). (The designated stoichiometry of the calcium silicate hydrate is only an approximation, since its composition is variable. It is usually referred to as C–S–H, which implies no definite stoichiometry or regular crystal structure.) The rate of hydration (Fig. 3) and the heat of hydration of the two silicates are quite different. The silicates are the components responsible for the strength of cement pastes, but in the first seven days the alite content controls the development of strength. The hydration of alite has an initial induction period that keeps concrete fluid for 1–2 hours and allows it to be handled and placed conveniently.

In the presence of gypsum, C_3A and the ferrite phase form calcium sulfoaluminate hydrates. The first product is the trisulfate form, ettringite, but when all gypsum has been consumed ettringite converts to the monosulfate form. Although C_3A and the ferrite phase contribute little to strength, they are still important components. When $C_4A\bar{S}H_{12}$ is exposed to an external source of sulfate ions, it will convert back to ettringite with concomitant expansion.

Interactions occur between the clinker compounds when they are hydrating in cement. For example, belite hydrates faster in the presence of alite, C_3A and the ferrite phase compete for the gypsum during their hydration, and gypsum accelerates the hydration of alite. During its formation, C–S–H can incorporate considerable quantities of alumina, sulfate, and Fe_2O_3 into solid solution. Thus less sulfoaluminates are formed during hydration than theory predicts; perhaps as little as half the expected amount.

It is important to know the effect of temperature on the hydration process because often concretes are cured at elevated temperatures. During the first 24 hours, the hydration of cement is very sensitive to temperature; apparent activation energies of $40\,kJ\,mol^{-1}$ have been reported for cement or alite. Thus, at low temperatures the rate of hydration, and hence of strength development, is very much reduced, while at high temperatures the reaction is accelerated.

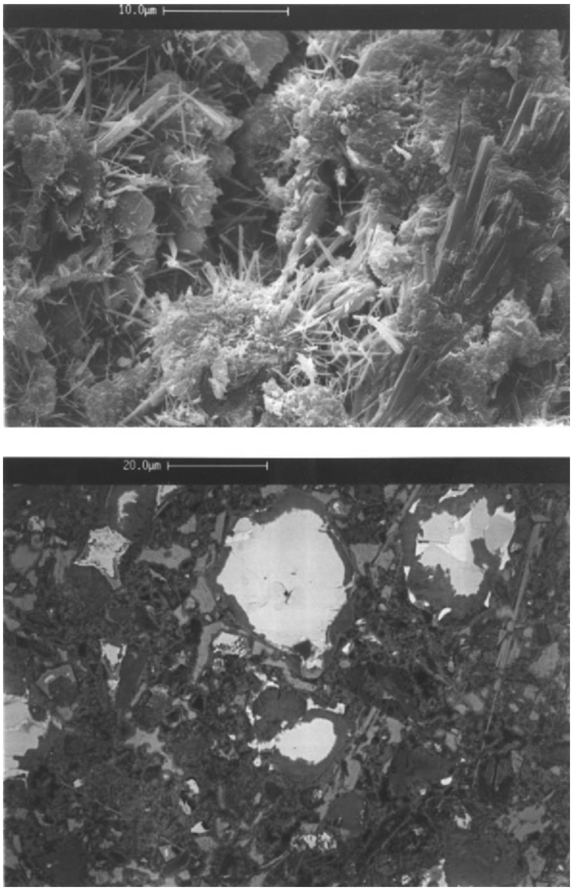


Figure 4
Microstructure of hardened cement paste revealed by scanning electron micrographs. (Courtesy of Dr. Paul Stutzman, National Institute of Standards and Technology).
Upper panel: fracture surface of a 7-day old paste. Ettringite needles are growing into waterfilled pores and monolithic $Ca(OH)_2$ on the right-hand side. (× 2500)
Lower panel: polished epoxy-impregnated 28-day old cement paste. Bright particles of unreacted cement are surrounded by dark grey C–S–H. The matrix between the particles is an intimate mixture of C–S–H (dark grey) and $Ca(OH)_2$ (light grey) (× 1000).

The calcium sulfoaluminate hydrates become unstable at 70–90 °C, while above 100 °C, C–S–H will convert to crystalline compounds. In the temperature range of autoclaved concrete products (150–200 °C under pressurized steam), C–S–H rapidly transforms to α-C_2SH. This results in low strength because the transformation increases porosity, but when finely divided silica is added, C–S–H converts to 1.1 nm tobermorite ($C_5S_6H_5$), which has good cementing properties.

Table 4
Composition of hardened paste (water–cement weight ratio, w/c = 0.5)

Volume component	Percent	Remarks
C–S–H	55	Continuous matrix: high intrinsic porosity; amorphous
CH	20	Large crystals (0.01–1 mm)
Calcium sulfoaluminate hydrates	10	Small crystals (1–10 μm)
Capillary porosity	~15	Depends on the amount of water used

Table 5
Influence of paste porosity on concrete properties.

Pore component	Size	Effect on properties
Capillary pores		
Macropores	> 0.1 μm (100 nm)	Permeability coefficient; Ionic diffusivities Strength (also total porosity)
Large mesopores	10–100 nm	Permeability coefficient; Ionic diffusivities Drying shrinkage
Gel pores		
Small mesopores	2.5–10 nm	Drying shrinkage; Creep
Micropores	< 2.5 nm	Drying shrinkage; Creep

4. Structure of Hardened Cement Paste

The transition of a hydrating cement paste from a fluid to a rigid material is a continuous process resulting from progressive hydration of the calcium silicates. Ettringite may control rheological properties, but does not form a continuous matrix because C_3A and the ferrite phase are minor components. Ettringite crystallizes mostly at the surface of the cement grains but also grows in the pore space (Fig. 4). Under special conditions crystallization of larger amounts of ettringite can cause premature stiffening because of physical interlocking of the needle-like crystals. The continuous matrix of hardened paste is formed as C–S–H coatings at the surface of calcium silicate particles merge as hydration proceeds. Calcium hydroxide grows in the water-filled voids between the hydrating particles. The approximate composition of the paste is given in Table 4.

C–S–H is amorphous colloidal material with high surface area and small particle size; these features can be seen only with an electron microscope. It has a large volume of intrinsic microporosity (~25 vol.%), which lowers its intrinsic strength and causes volume instability on drying. To a large extent, the material properties of concrete reflect the properties of C–S–H. In contrast, massive crystals of calcium hydroxide can be readily seen, even by optical microscopy. Calcium hydroxide contributes to strength because it fills voids, but it is highly soluble. The presence of soluble alkali salts makes the concrete pore solution very alkaline, and this can reduce the durability of concrete in certain circumstances. The calcium sulfoaluminate hydrates play a relatively minor role in

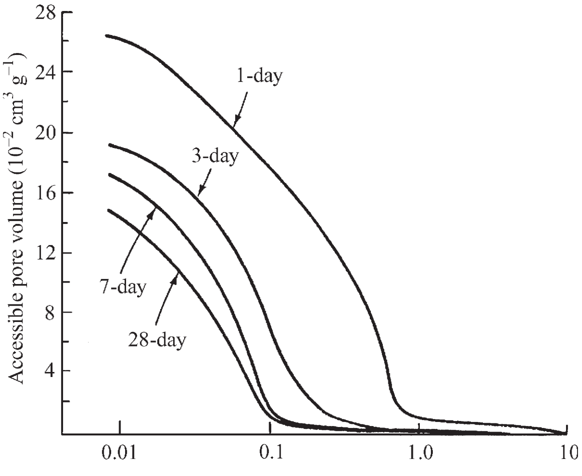


Figure 5
Accessibility of the pore volume of hydrated cement paste as measured by mercury intrusion porosimetry.

Table 6
Mineral admixtures used in concrete.

Admixture	Form of reactive silica	Reaction
Fly ash– Class F	Aluminosilicate glass	Pozzolanic
Class F	Calcium aluminosilicate glass	Pozzolanic (cementitious)
Silica fume	Amorphous silica ($\sim 20\,m^2\,g^{-1}$)	Pozzolanic
Granulated blast furnace slag	Calcium aluminiosilicate glass	Cementitious + pozzolanic
Calcined clay	Amorphous silica (and alumina) ($\sim 10\,m^2\,g^{-1}$)	Pozzolanic
Rice husk ash	Amorphous silica ($\sim 60\,m^2\,g^{-1}$)	Pozzolanic

the behavior of hardened pastes, except when sulfates are present.

All properties of concrete are affected by the volume of capillary pores in the hardened cement paste. The pores are interconnected, allowing water and dissolved species to penetrate the concrete, with potential adverse consequences for service life (i.e., durability). Transport properties such as water permeability and ionic diffusivity are controlled by the size and volume of pores. A pore size distribution can be obtained using mercury intrusion porosimetry (Fig. 5). The volume of pore space accessible through a pore diameter $>0.1\,\mu m$, and the pore diameter corresponding to the maximum rate of intrusion are the main parameters controlling transport. The capillary pore volume is controlled by the water–cement ratio and the extent of hydration during curing. Capillary pores also affect other properties as well (see Table 5).

4.1 Modification of Microstructure

The microstructure can also be modified by the use of reactive aluminosilicate (mineral admixtures), such as fly ash, silica fume, or granulated blast furnace slag. These compounds contain amorphous silica and react chemically to convert calcium hydroxide to C–S–H (see Table 6). This the so-called *pozzolanic reaction*. In addition, slag will react cementitiously to form C–S–H directly. The C–S–H formed by mineral admixtures has a lower C/S molar ratio than that formed by direct hydration.

Increasing the volume factor of C–S–H in the microstructure reduces the overall porosity because C–S–H has a lower specific gravity than $Ca(OH)_2$. The proportion of micropores and small mesopores also increases. As a result such pastes have enhanced resistance to the transport of water and dissolved species, thereby creating more durable concrete. The reduction of the volume fraction of leachable $Ca(OH)_2$ and of the pore solution also enhances durability.

5. Concluding Remarks

Portland cement is a calcium silicate cement formed by high temperature reactions between limestone and clay. The reaction of calcium silicates with water (hydration) are responsible for the strength and durability of concrete. Small quantities of alumina and iron oxide which act as fluxing agents during manufacture form calcium aluminates, which also react with water. The hydration products are arranged physically to form a solid structure with a network of small pores. Hydration chemistry and the resulting microstructure can be modified by the addition of reactive aluminosilicates.

Bibliography

Barnes P (ed.) 1984 *Structure and Performance of Cement*. Applied Science, Barking, UK

Bye G C 1983 *Portland Cement: Composition, Production and Properties*. Pergamon, Oxford

Francis A J 1977 *The Cement Industry 1796–1914: A History*. David and Charles, Newton Abbot, UK

Ghosh S N (ed.) 1983 *Advances in Cement Technology*. Pergamon, Oxford

Kurdowski W 1991 In: Ghosh S N (ed.) *Cement and Concrete Science and Technology*. Vol. 1, Part 1, ABI Books, New Delhi, India, pp. 58–98 and pp. 201–36

Lea F M, Hewlett P C 1998 *Lea's Chemistry of Cement and Concrete*, 4th edn. Arnold, London

Mehta P K, Montiero P J M 1993 *Concrete: Microstructure, Properties, and Materials*. McGraw-Hill, New York

Mindess S, Young J F 1981 *Concrete*. Prentice Hall, Englewood Cliff, New Jersey, USA

Skalny J P (ed.) 1989 *Materials Science of Concrete. Vol. I*. American Ceramic Society, Westerville, OH, USA

Taylor H F W 1997 *Cement Chemistry*, 2nd edn. Thomas Telford, London

J. F. Young
University of Illinois, Urbana,
Illinois, USA

Powders, Solution Synthesis of

The purpose of this article is to describe solution-based technologies for controlling powder characteristics of a spectrum of oxide and nonoxide ceramic materials. It will also highlight the advantages and drawbacks that each solution-based method offers in terms of the powder characteristics that can and cannot be controlled. For further information on vapor phase methods see *Vapor Synthesis*#.

Ceramics, whether they are single phase or composite materials, are typically based on multicomponent chemistries where more than one cation and/or anion must be incorporated, e.g., barium titanate and silicon oxynitride. For the processing of typical polycrystalline ceramics, desirable physicochemical characteristics of a ceramic powder include submicron particle size, controlled particle size distribution, uniform and equiaxed shape, minimum degree of agglomeration, high degree of chemical purity, controlled chemical and phase homogeneity, low intraparticle density, minimum weight loss upon heating, and controlled surface chemistry.

Conventional powder synthesis technology, discussed in detail in *Calcination and Phase Transformations*#, requires the three consecutive steps of mixing, solid-state reaction, and milling. This overall process produces undesirable characteristics, which include phase and stoichiometric heterogeneities and poor control of morphology, particle size distribution, and aggregate population (chemically bound agglomerates). In many cases, remedial processing is required to improve the characteristics of conventional powders, such as repeated calcination and milling steps. The limitations of conventional technology drive the research, development, and commercialization of chemical (solution- and vapor-phase) synthesis methods for making powders.

Solution-based powder synthesis technologies utilize a broad base of technologies. The technologies discussed in this article all use a liquid phase to solubilize one or more components in the process. These technologies will be categorized by identifying where the chemical supersaturation takes place that produces the particulate phase. This review encompasses technologies that use chemistry in the bulk liquid phase, chemistry between heterogeneous phases, and chemistry in a droplet.

1. Chemistry in the Bulk Liquid Phase

Chemistry in the bulk liquid phase is an approach in which the physical characteristics of a ceramic powder are controlled by nucleation, growth, and aging processes in solution. In these systems, precipitation is driven by chemical supersaturation using one of the following synthetic methods: direct strike precipitation, solvent removal, nonsolvent addition, or precipitation from homogeneous solution.

1.1 Direct Strike Precipitation

Direct strike precipitation, commonly employed for a range of materials, involves the addition of a precipitating agent to a solution of soluble metal salts to form a sparingly soluble salt. When the soluble salts are added to a solution containing the precipitating agent, the procedure is referred to as a reverse strike process.

In direct strike precipitation, copious precipitation ensues at a rate faster than the timescale required for mixing of the reagents. For example, in the preparation of lanthanum-modified lead zirconate titanate, lead and lanthanum nitrates, zirconium oxychloride, titanium tetrachloride, and hydrogen peroxide are dissolved in water at 40 °C. Precipitation is induced by the addition of ammonium hydroxide to this solution to maintain a pH of ~9. The precipitate, most likely a mixture of various hydroxides, is filtered from suspension, washed free of soluble species with hot distilled water and dried at 100 °C. Calcination at a temperature of 650 °C is necessary to form the oxide, followed by subsequent milling to break up aggregates formed during precipitation and calcination.

Nonoxides can be prepared by direct strike precipitation. Aluminum hydride can be reacted with ammonia at room temperature in tetrahydrofuran to yield a precipitate ($Al(NH)NH_2$) that can be calcined at 1100 °C to form aluminum nitride. Silicon diimide ($[-Si(NH)_2-]_x$) can be prepared as follows:

$$SiCl_4 + 18NH_3 \rightarrow [-Si(NH)_2-]_x + 4NH_4Cl \cdot 3NH_3 \quad (1)$$

Submicron silicon nitride (95% α-Si_3N_4/5% β-Si_3N_4) has been prepared by pyrolysis of silicon diimide at temperatures of 1200–1700 °C. Ammonium salts that coprecipitate with the desired product must be eliminated via washing prior to pyrolysis of the precipitate.

The common ion effect can cause solution homogeneity problems in multicomponent precursor solutions when one of the soluble salts has an anion that renders one or more of the cation components insoluble. Such a heterogeneous precursor system is unlikely to produce a physically and chemically uniform powder. Thus, the solubility of all salts that can form in the system should be examined to avoid this problem.

Differential solubilities of the various precipitating phases also affect the precipitation kinetics of each component, causing the stoichiometry of the precipitate to vary with reaction time. The copious nature of a direct strike precipitation creates concentration gradients in solution that can further accentuate this effect. Differential solubility problems can be minimized by introducing a mixture of anions that renders all cations highly insoluble. For instance, manganese–zinc ferrite can be prepared by addition of ammonium hydroxide (NH_4OH) and ammonium bicarbonate (NH_4HCO_3) to metal sulfate solutions to prepare a hydroxide–carbonate coprecipitate.

To minimize segregation during direct strike precipitation and to optimize control of the stoichiometry, multicomponent salts of fixed stoichiometry can be precipitated. $BaTiO_3$ with a Ba:Ti molar ratio of 0.995 ± 0.005 has been prepared by calcination of $BaTi(C_6H_6O_7)_3 \cdot 6H_2O$. In order to prepare this salt, titanium tetrabutoxide, citric acid, and barium carbonate are mixed in aqueous solution in a pH range of 0.5 to 2.6. At a pH of 3.2, a $Ba_2Ti(C_6H_6O_7)_2 \cdot 7H_2O$ salt forms instead.

Other undesirable effects such as coprecipitated anion impurities (e.g., chlorides and sulfates) plague direct strike precipitation. These impurities can be incorporated into the particles as lattice defects or can adsorb onto powder surfaces. In addition, soluble species in solution can become occluded within agglomerates of particles.

Proper choice of the soluble anions in solution is important in order to minimize the above problems. Anions such as nitrates, which decompose at low temperatures, are more desirable since they can be eliminated from the powder during calcination. Anion contamination problems can be avoided with the use of "anion-less" precursors such as metal alkoxides. Metal alkoxides react with water to precipitate ceramic polymers by hydrolysis:

$$M(OR)_4 + H_2O \rightarrow (OR)_3M{-}OH + ROH \tag{2}$$

and polymerization:

$$(RO)_3M{-}OH + HO{-}M(OR)_3 \rightarrow (RO)_3M{-}O{-}M(OR)_3 + H_2O \tag{3}$$

$$(RO)_3M{-}OH + RO{-}M(OR)_3 \rightarrow (RO)_3M{-}O{-}M(OR)_3 + ROH \tag{4}$$

In many cases, a nearly fully cross-linked oxide powder (e.g., silica) can result, subsequently forming a hydrous oxide. However, the formation of various types of hydrates, such as $AlO(OH)$ or $Sr(OH)_2$ that cannot fully or even partially polymerize according to Eqns. (3) and (4), is also possible.

Metal alkoxides are useful for preparing uniform multicomponent solutions since most exhibit excellent solubility in organic solvents such as alcohol, toluene, or tetrahydrofuran. Multicomponent alkoxide formation can enhance solution solubility further. For instance, solutions of $MgAl_2(O\text{-}i\text{-}Pr)_8$ can be prepared from reactions between magnesium and aluminum alkoxides. The complex can be hydrolytically decomposed to form a coprecipitate of hydroxides and calcined at 800 °C to prepare submicron spinel ($MgAl_2O_4$). When a multiphase coprecipitate is undesirable, condensation complexes can be prepared by partial hydrolysis routes:

$$M(OR)_4 + H_2O \rightarrow (RO)_3M{-}OH + ROH \tag{5}$$

$$(RO)_3M{-}OH + RO{-}M'(OR)_3 \rightarrow (RO)_3M{-}O{-}M'(OR)_3 + ROH \tag{6}$$

This approach has been utilized in the preparation of multicomponent ceramic powders such as cordierite ($2MgO \cdot 2Al_2O_3 \cdot 5SiO_2$) with no evidence of segregation.

Alkoxides can be complexed with other metal organic compounds or inorganic salts to prepare useful precursors for powders. For instance, barium hydroxide can be reacted with titanium isopropoxide in methanol to prepare a soluble complex as a precursor for barium titanate, and lead acetate and titanium isopropoxide can be reacted in methoxymethanol as a precursor for lead titanate.

1.2 Solvent Removal and Nonsolvent Addition

Sol–gel methods are generally useful for preparing powders since the gels can be broken up into granules during the drying step or ground in the dry state to produce granules of high surface area material. Supercritical CO_2 extraction, in which an overpressure of supercritical CO_2 is applied to extract the solvent from the gel, is another route for powder preparation. A constant flow of supercritical CO_2 is maintained until alcohol can no longer be seen separating from the CO_2 leaving the autoclave. Solvent is removed from the autoclave by depressurizing the gel at the rate required to maintain a temperature above the critical temperature of the fluid while drying the gel. Highly microporous single-component (silica, titania, etc.) and multicomponent powders (aluminum titanates, aluminosilicates, etc.) have been prepared with surface areas as high as $950\,m^2\,g^{-1}$ and disbursed particle sizes of less than 1 μm.

In addition to solvent removal techniques, there are also direct precipitation methods based on addition of a nonsolvent. In this case, an aqueous solution is prepared and precipitation is induced by dilution of this solution with a hydrophilic nonsolvent. Soluble salts employed include sulfates, citrates, and nitrates, depending on the assortment of cations desired. Alcohol (ethanol, isopropanol, etc.) or acetone typically is employed as the nonsolvent. Proper choice of the salt concentration, solution environment (pH, addition of complexing agents) and nonsolvent is essential to achieve the proper stoichiometry and satisfactory homogeneity. Materials such as barium titanate, lead zirconate, magnesium aluminate, and manganese ferrites can be prepared. This method and other approaches such as freeze drying and evaporation are commonly employed in chemistry in a droplet approach, discussed later in this article.

1.3 Drawbacks of the Above Technologies

In most liquid-phase approaches (including all of the cases described above), the precipitation product must be calcined to form the oxide and thus is subjected to the same types of problems inherent in conventional processing. Reduced particle size and greater homogeneity accelerate the reaction kinetics, as mentioned earlier, resulting in lower calcination temperatures. However, the precipitate also becomes a great deal more aggregated (Fig. 1). Thus, in many cases, a milling step is required, allowing impurities to be introduced into the powder.

In a few cases, direct strike precipitation can result in the direct formation of oxides. Titanium alkoxides can be hydrolyzed with an aqueous solution of barium hydroxide, producing submicron agglomerates of $\sim 10\,\mathrm{nm}\,\mathrm{BaTiO_3}$. The same result can be accomplished by hydrolyzing a barium–titanium double alkoxide dissolved in benzene or toluene. The use of anions such as tungstate, chromate, or molybdate is useful for the preparation of oxides directly from solution. For instance, barium chromate can be precipitated from a solution of potassium chromate and barium chloride. Oxides also can be precipitated directly from solution via oxidation reactions. Basic aqueous solutions of iron nitrates or sulfates can be reacted with oxygen to prepare polymorphs of Fe_2O_3 and other multicomponent ferrites (Mn, Zn, or Co) at temperatures as low as 70 °C.

Despite advances in direct strike precipitation technology, the inability to avoid nonuniformities in solution hinders control of the physical characteristics of the powder. For example, high particle number densities and high electrolyte concentrations encountered during direct strike precipitation lead to a low degree of colloid stability and a concomitant high degree of agglomeration. As growth and aging proceed, aggregates (hard agglomerates) are formed. The presence of dispersants during precipitation can reduce aggregation. Polyethylene glycol has been used as an *in situ* dispersant for the preparation of 10 nm yttrium-stabilized zirconia.

1.4 Precipitation from Homogeneous Solution

Precise control of physical powder characteristics requires uniform supersaturation in solution. When this is achieved, nucleation precedes growth and aging, and these processes occur uniformly throughout the reaction medium. Controlled precipitation is best accomplished using a precipitation from homogeneous solution (PFHS) technique. This approach utilizes chemical reactions whose kinetics somehow limit the rate of release of supersaturating species.

In order to control particle size, size distribution, and morphology, special attention must be given to concentration, pH, heating rate, temperature, corresponding anion of the soluble metal salt, choice of precipitating reaction, and aging conditions. For monodisperse powders (Fig. 2), precipitation must occur at concentrations typically less than 0.05 M and at a pH near the solubility limit of the precipitate. At higher concentrations morphology nonuniformity and aggregation result.

A detailed understanding of the interplay between solution processing variables and physical powder characteristics is lacking. Thus, the appropriate precipitation conditions for each system must be investigated empirically. For instance, the use of chloride vs. nitrate solutions to precipitate α-Fe_2O_3 under otherwise identical conditions yields elliptical vs. star-shaped particles.

The most commonly used PFHS technique uses a soluble salt codissolved with a precursor that decomposes upon heating, releasing an anion that quantitatively precipitates the dissolved cation. Zinc oxalate dihydrate also can be precipitated from an aqueous oxalic acid–zinc nitrate solution via anion release by hydrolyzing diethyl oxalate in 15 vol.% water in a nonaqueous acetic acid solution:

$$\mathrm{Et_2C_2O_4} + 2\mathrm{H_2O} \rightarrow 2\mathrm{EtOH} + 2\mathrm{H^+} + [\mathrm{C_2O_4}]^{2-} \quad (7)$$

The release of the oxalate dianion is being regulated by the rate of diethyl oxalate decomposition in a solution of dissociated Zn^{2+} to precipitate $ZnC_2O_4 \cdot 2H_2O$ as shown below:

$$\mathrm{Zn^{2+}} + [\mathrm{C_2O_4}]^{2-} + 2\mathrm{H_2O} \rightarrow \mathrm{ZnC_2O_4} \cdot 2\mathrm{H_2O(s)} \quad (8)$$

This PFHS system precipitates $\sim 10\,\mu\mathrm{m}$ (~ 7 aspect ratio) needle-like $ZnC_2O_4 \cdot 2H_2O$ particles.

PFHS techniques can also be used to prepare multicomponent powders with controlled stoichiometry and phase homogeneity. In order to obtain a phase-homogeneous powder with the intended cation stoichiometry, careful attention must be paid to preparation of the homogeneous solution and the types of multicomponent complexes that form. Problems such as stoichiometry control that are associated with multicomponent monodisperse syntheses can be avoided when multilayer structures are prepared, since each component is added in a separate processing step. A magnetic material was prepared by precipitating Fe_2O_3-based compositions via oxidation of Fe^{2+} solutions on a porous silica carrier in suspension.

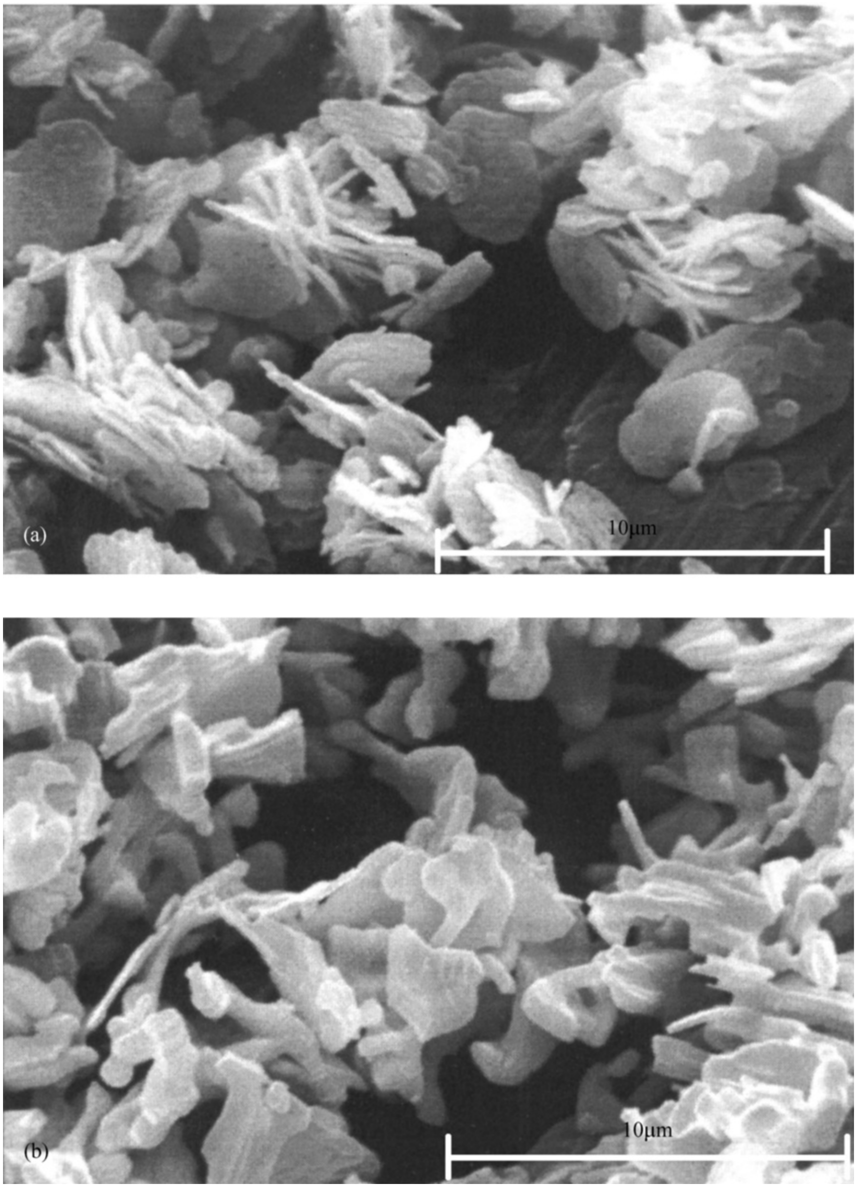


Figure 1
Lead oxalate: (a) as-precipitated from solution; (b) calcined at 600 °C in air for 2 h to form lead oxide.

In nonaqueous media, controlled hydrolysis of metal alkoxides with dilute alcoholic solutions provides the simplest means of preparing spherical monodisperse powders and composite powders. At times, alcohol solvents produce inorganic gels instead of precipitate, necessitating the use of other water-soluble solvents such as acetonitrile. Single and multicomponent oxides such as hydrous titania, hydrous silica, strontium–titanium hydrous oxides, and yttria-doped zirconium hydrous oxide can be prepared by hydrolysis of metal alkoxide solutions. Composite particles such as alumina coated with hydrous titania have been prepared for alumina–titania composites. *In situ* dispersants such as hydroxyl propyl cellulose have been used in nonaqueous PFHS systems for monodisperse titanium hydrous oxide.

Nonoxide powder preparation is possible via a small number of PFHS techniques. Submicron diborides of titanium, zirconium, and hafnium are prepared by refluxing metal borohydrides in

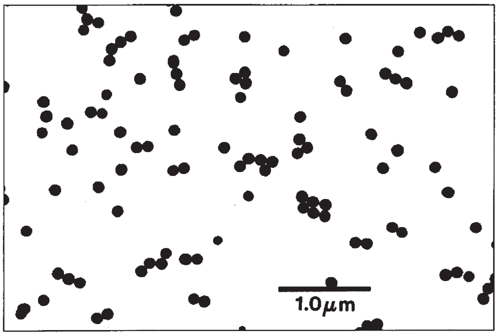


Figure 2
Monodisperse lead hydroxycarbonate prepared by a precipitation from homogeneous solution method.

xylene at 140 °C:

$$Zr(BH_4)_4 \rightarrow ZrB_2 + B_2H_6 + 5H_2 \quad (9)$$

The powders are amorphous as precipitated but crystallized when annealed. $SiCl_4$ was dehalogenated and carburized at 130 °C in *n*-heptane to form 1–5 μm ($70\,m^2\,g^{-1}$) amorphous silicon carbide:

$$SiCl_4 + CCl_4 + 8Na \rightarrow SiC + 8NaCl. \quad (10)$$

By using the chloride pairs $TiCl_4$ and BCl_3, $TiCl_4$ and CCl_4, or BCl_3 and CCl_4, the nonoxides TiB_2, TiC or B_4C, respectively, were obtained.

2. *Chemistry Between Heterogeneous Phases*

Chemistry between heterogeneous phases is an approach in which the physical powder characteristics are controlled by interaction of a discrete solid phase with a continuous one. Synthetic methods falling into this category include hydrothermal synthesis, molten salt synthesis, pyrolysis, and spark erosion.

2.1 Hydrothermal Synthesis

Hydrothermal synthesis utilizes solutions, gels, and/or suspensions as precursors to precipitate anhydrous oxides in a single step at elevated pressures (e.g., ~14–160 atm) and metric temperatures (e.g., 100–350 °C). The low reaction temperatures, relative to those of synthesis methods based on solid-state reaction, are due to a solution-phase mass transport instead of solid-state diffusion. The combined effect of pressure and temperature can also reduce free energies for various equilibria. This serves to stabilize phases that might not be stable at atmospheric conditions. There are many instances where anhydrous phases are unstable in aqueous solutions. In these cases, solvothermal synthesis is employed where nonaqueous liquids such as 1,4-butanediol, ammonia, or isopropanol are used as alternative precipitation media. In addition, efforts are under way to develop processes that operate above the critical point of water (374 °C, 22.4 MPa); however, reactor safety and cost present significant technological barriers.

Hydrothermal reactions are sometimes referred to as high-temperature PFHS systems since uniform solutions of aqueous complexes prepared under atmospheric conditions are reacted at elevated temperature and pressure instead. For instance, 200 °C thermal decomposition of calcium EDTA (ethylenediaminetetraacetate) complexes in pH 7.5 phosphate solution has been used to prepare 100 μm hexagonal rods of carbonate-substituted hydroxyl apatite.

Ceramic powders can be prepared from suspensions or gels via an *in situ* transformation and/or a dissolution–reprecipitation. The former can be effected in a dispersion when suspended particles undergo a polymorphic or chemical phase transformation. The control of particle morphology and size is a function of the characteristics of the suspended solid phase precursor. For instance, narrow size distribution spherical submicron titanium hydrous oxide can be chemically transformed into polycrystalline anhydrous anatase spheres (TiO_2) of the same size by refluxing the powder for 12 h in a pH 10 ammonium hydroxide–water suspension at atmospheric pressure and a temperature of 100 °C.

Dissolution–reprecipitation of suspensions or gels can be effected when suspended particles dissolve into solution, supersaturate the solution phase, and subsequently precipitate particles (Fig. 3). In many cases, the suspended solids are not very soluble in water, and solubilizing components (mineralizers) such as acids, bases, or other complex-forming substances must be added. For example, in the preparation of zirconia powder from a suspension of hydrous zirconium oxide, a hydrothermal reaction at low pH (<3.5) dissolves the zirconium hydrous oxide precursor in soluble species that subsequently reprecipitate the oxide. Additives play a significant role in controlling particle size, morphology, and chemical reaction rate in a manner very similar to that observed in PFHS systems. For instance, a fraction of a pH unit change in the hydrothermal solution determines whether $PbTiO_3$ forms as a platelet- or cube-shaped particle.

Hydrothermal oxidation of metals is another interesting route for preparing ceramic powders. Metals such as iron, zirconium, aluminum, or hafnium can be reacted with water at temperatures as low as 500 °C and 100 MPa:

$$mM + nH_2O \rightarrow M_mO_n + nH_2 \quad (11)$$

Multicomponent oxides such as alumina–zirconia composite powders can be prepared from a Zr–Al

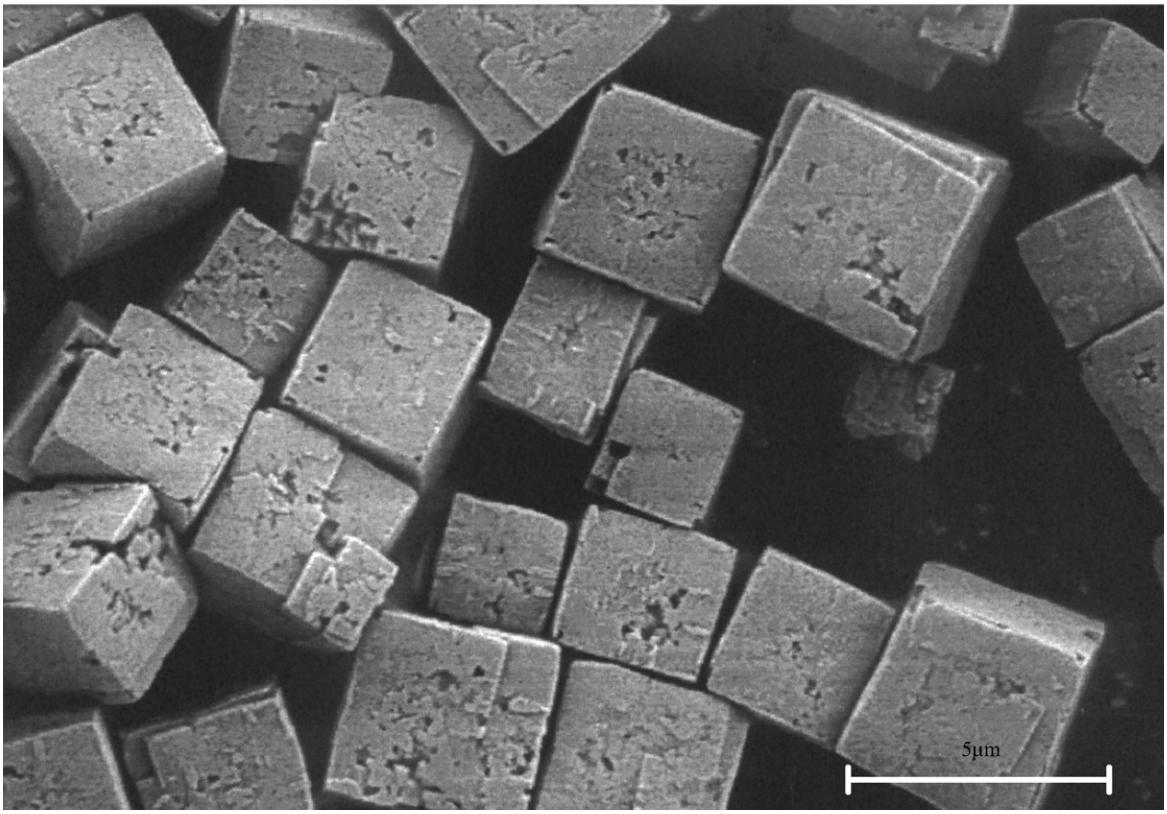


Figure 3
Cube-shaped lead zirconate titanate prepared via hydrothermal methods.

metal alloy. Metals such as titanium or zirconium can be reacted with ammonium chloride in an ammoniacal medium (rather than water) to form nitrides at rather high temperatures (500–600 °C) and pressures (1 MPa).

2.2 *Molten Salt Synthesis*

Molten (fused) salt synthesis is similar in many ways to a dissolution–reprecipitation hydrothermal synthesis. An inert molten ceramic material in which the ceramic powders are somewhat soluble acts as a solvent medium. The single-component oxides dissolve, react in solution to form multicomponent species and precipitate from solution. For example, lead zirconate titanate can be prepared by first mixing together TiO_2, ZrO_2, PbO, and any desired dopants to form an aqueous slurry. A KCl–NaCl molten salt eutectic is mixed into the slurry, causing the system to gel. Upon drying, the mixture is heated to ~980 °C for 1 h to form the lead zirconate titanate, which subsequently is rinsed free of the molten salts and any excess unreacted ions. In some systems, the solvent can react with the other components to yield multicomponent powders:

$$TiO_2 + 2MgCl_2 + 2SiO_2 \rightarrow TiCl_4 + 2MgSiO_3 \qquad (12)$$

Since $TiCl_4$ is highly volatile, purification of the precipitated powder is easily accomplished. Materials such as spinel, cordierite, and magnesium titanate have been prepared in this manner.

Particle formation in molten salt synthesis has been observed to depend upon the dissolution characteristics of the individual components. When a slowly dissolving component is in the presence of a rapidly dissolving one, the multicomponent reaction product precipitates on the surface of the less soluble component. Unidirectional diffusion of the more soluble component through the product layer carries the process to completion, resulting in a powder with a morphology reflective of the more slowly dissolving component. Both needle-like barium titanate and plate-like zinc ferrite have been prepared in this manner. On the other hand, feedstocks with somewhat equivalent dissolution characteristics precipitate

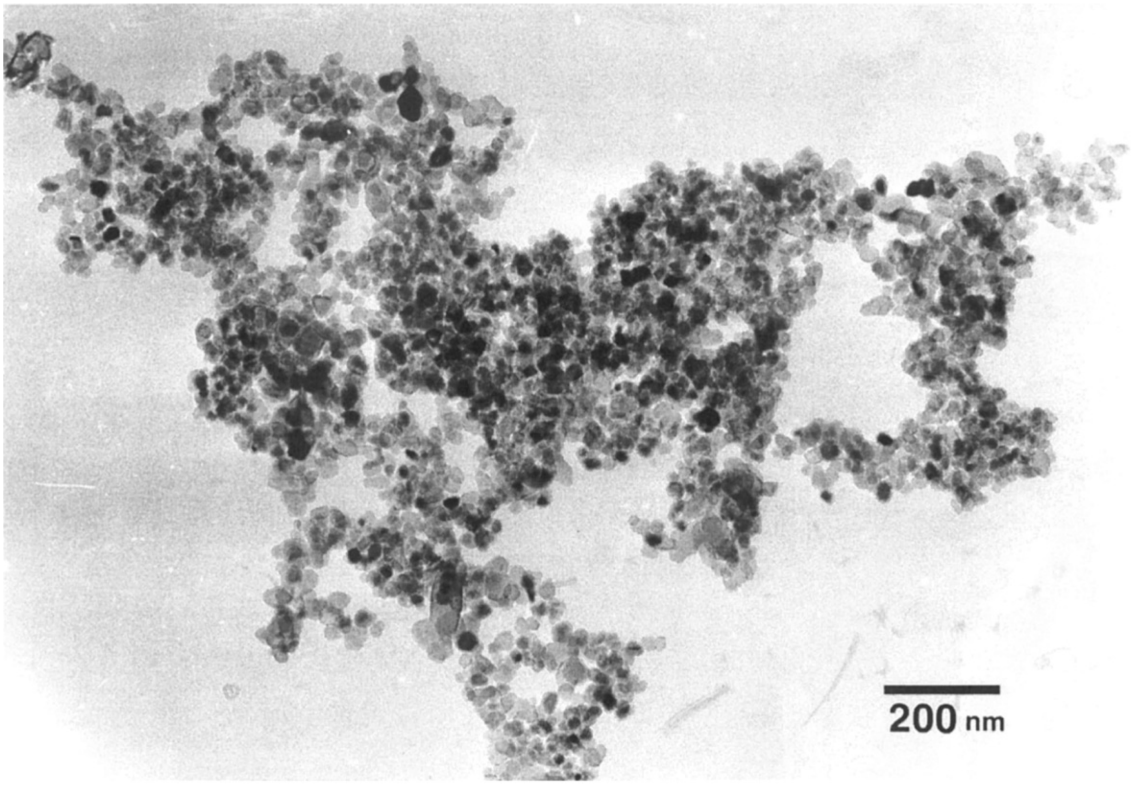


Figure 4
Nanostructured aluminum nitride prepared by pyrolyzing aluminum-based Pechini polymers in nitrogen at 1500 °C for 4 h.

in bulk solution, producing more equiaxed morphologies. Oxides such as Bi_2WO_6 and $PbNb_2O_6$ are believed to form in this manner.

2.3 *Pyrolysis*

Pyrolysis processes form a ceramic powder by solid-state decomposition of a solid or molten precursor (typically a molecular or polymeric compound) in the presence of an inert (e.g., vacuum, argon) or reactive (e.g., oxygen, nitrogen) atmosphere. This method is distinguishable from a solution precipitation/calcination method because it is the pyrolysis rather than the chemical synthesis that influences the physical characteristics of the powder. For instance, agglomeration and excessive crystallite growth can be caused by excessive exothermic decomposition. Decomposition residue (e.g., carbon) can result from factors such as poor choice of pyrolysis atmosphere and control, improper gas flow rates, and furnace choice. Mechanistic studies of the decomposition and its interaction with the processing variables are essential for control of the powder characteristics of the pyrolysis product.

Metal–organic molecular compounds such as carboxylates are useful materials for preparing oxides and nonoxides. These compounds can be derived in-house from starting reagents such as metal alkoxides, nitrates, or chlorides using simple procedures. Neodecanoate salts of various metals form uniform solutions in xylene. These solutions can be pyrolyzed to form submicron powders. For instance, $PbTiO_3$ can be prepared from a solution of dimethoxytitanium dineodecanoate ($(\mathrm{(MeO)_2Ti(OCOC_9H_{19})_2})$) and lead dineodecanoate ($\mathrm{Pb(OCOC_9H_{19})_2}$) by pyrolyzing at temperatures of ~600 °C. Molecular metal–organic precursors are useful for the preparation of nonoxide ceramics such as boron nitride using nitrogen-bearing molecular precursors such as $\mathrm{B(N(C_2H_4O)_3)_3}$. Pyrolysis in a nitrogen-bearing atmosphere such as nitrogen or ammonia is critical. In general, most pyrolysis processes can be modified to produce nonoxides by paying careful attention to the pyrolysis atmosphere. Thermodynamic approaches used for engineering conventional solid state reactions, such as halogenation for halides, carbothermal nitridation for nitrides, and carbothermal reduction for carbides, provide the necessary insight for proper choice of atmosphere and firing schedule to make the chosen nonoxide material.

Choice of ligand in the molecular precursor is an important processing variable since it controls pyrolysis behavior. Crystalline solids such as the hydrazinium metal hydrazinecarboxylate hydrates, easily prepared from aqueous metal sulfate solutions, autocatalytically decompose at low temperatures to form oxide powders. For example, a compound such as $N_2H_5Mn_{1/3}Fe_{2/3}(OCONHNH_2)_3 \cdot H_2O$ decomposes in a single step in a temperature range of 75–120 °C to form agglomerates of nanometer-sized $MnFe_2O_4$. Inorganic nitrates (oxidizers) can be combined with organics such as urea (fuel) in stoichiometric compositions using principles of propellant chemistry to produce 4–5 μm agglomerates of submicron particles at temperatures as low as 500 °C. A diverse number of oxides, including spinel ($MgAl_2O_4$), $Y_3Al_5O_{12}$ (YAG), ZrO_2–Al_2O_3 composites and α-Al_2O_3, have been prepared with this method.

Polymeric metal organic precursors prepared from water-soluble metallic salts of polyfunctional acids are useful for the preparation of multicomponent oxides. Typically, α-hydroxy acids (e.g., hydroxybutyric acid, citric acid) and/or α,β-unsaturated acids (maleic acid, acrylic acid) are mixed with metal salts to form a uniform solution of metal cation complexes. The solution subsequently is evaporated until a viscous mass is obtained, followed by a drying step at $\sim$120 °C. During the dissolution and solvent removal stages, it is essential that none of the components crystallizes, so that powder homogeneity is ensured. Finally, the precursor powder is calcined ($>$600 °C). Materials such as $BaFe_{12}O_{19}$, $LaCrO_3$, and $Y_3Fe_{4.9}Al_{0.1}O_{12}$ have been prepared with this method. A variant of the above process is the liquid mix process, where a polyfunctional alcohol such as ethylene glycol is added to the acid to make a polyester polymeric residue that can complex practically any metal ion in the periodic table. Many ceramic compositions, including aluminates ($LaAlO_3$), silicates (Zn_2SiO_4), zirconates ($CaZrO_3$), and titanates ($Pb(Zr,Ti)O_3$), and nitrides (AlN) have been prepared by this method.

Organometallic polymers are considered desirable for nonoxide materials free of oxygen contamination since the metal center is bonded to carbon rather than oxygen. However, the synthesis of organometallic compounds requires air-sensitive procedures that can be complex. Many problems must be overcome with this approach, such as obtaining the desired phase assemblage, controlling chemical purity, diminishing the loss of volatile species that belong in the final ceramic product (e.g., Si, C, or N), minimizing the weight loss of organic species, and controlling the physical characteristics of the powder. Various nonoxide powders such as silicon nitride and silicon carbide have been prepared by decomposing polymeric precursors, sometimes producing a powder and other times yielding a porous monolithic form. The decomposition of the materials to produce a desired phase requires a combination of skill and art. For instance, silazane polymer obtained by KH treatment of $[MeSiHNH]_x$ cyclics in diethyl ether pyrolyzes in nitrogen to form α-silicon nitride at 1420 °C. When the synthesis is conducted in tetrahydrofuran, silicon nitride/silicon carbide composite powders form instead. Composite powder mixtures of silicon nitride and aluminum nitride can be made by pyrolyzing a mixture of polycarbosilane and diisobutylaluminum amide.

2.4 *Spark Erosion*

Another method that uses reactions between heterogeneous phases for preparing submicron ceramic powders is spark erosion. In this technique two electrodes are immersed in a dielectric liquid (organic or cryogenic) with a charge of large granules of the electrode material between the electrodes. A field strength of $\sim 10^5\,\mathrm{V\,cm^{-1}}$ is applied in a pulsed mode (230–610 kHz) to generate a spark discharge that takes place at the asperities of the electrode and granular charge. Since the material is heated in localized regions well above its melting point via plasma channel transient pressure zones, molten droplets or vapor species are produced. When the species leave the localized region they either freeze or condense to form particles. Subsequent reaction between the dielectric fluid or decomposition products and the powder (either as molten droplets or vapor) produces a ceramic powder. Spherical narrow-size distribution $\sim$34 μm titanium carbide has been prepared in a dodecane fluid using a titanium metal electrode, a graphite electrode, and a granular charge of titanium metal. Spark erosion of aluminum electrodes and charge in water produces submicron acicular aluminum hydroxide. This product is considered to form from condensed aluminum vapor species instead of molten droplets.

3. *Chemistry in a Droplet*

Chemistry in a droplet is an approach in which one of the earlier described synthetic methods occurs within the confines of a solvent droplet. The droplet not only controls the physical powder characteristics but also limits chemical segregation to the scale of the droplet diameter. Synthetic methods falling into this category involve the use of either liquid phase systems such as emulsions or microemulsions or liquid–gas phase systems such as aerosols.

3.1 *Emulsions*

An emulsion is a mixture of two immiscible liquids in which one liquid is dispersed in the other with the aid of a surfactant. The surfactant stabilizes the emulsion by reducing interfacial tension, reinforcing the interfacial film, and preventing coalescence of the droplets. Mixing conditions and surfactant variables control the size of the droplet. High shear rates and surfactant levels promote smaller droplet sizes, which can range from one to 50 μm. A single surfactant such as sorbitan monooleate can be employed for water-in-2-ethylhexanol emulsions. In some cases, a mixture of surfactants is employed. Sorbitan monostearate and polyoxyethylene sorbitan monolaurate mixtures are useful for water-in-toluene emulsions.

The water-in-oil emulsion has been the most widely used type of emulsion for the preparation of single-component oxides such as titania, multicomponent oxides as complex as synroc (a mixture of $BaAl_2Ti_6O_{16}$, $CaZrTi_2O_7$, and $CaTiO_3$), and nonoxides such as the carbides of thorium or uranium. In this type of emulsion, a stable colloidal sol consisting of one or more metal complexes is prepared and emulsified in an oil medium to form spherical droplets. Gelation of the spheres is induced in one of two ways. First, a liquid removal technique such as dehydration or evaporation can be used. For example, large volumes of an organic phase such as a long-chain alcohol (2-ethylhexanol) can be used to concentrate the sol by dehydrating the droplets. Upon sufficient dehydration, van der Waals forces predominate and flocculate the sol within the droplet. It is conceivable that a precipitation reaction occurs inside the droplet as well. Evaporation methods such as the "hot kerosene method," in which either an emulsion is added to a hot oil phase or the water is distilled from the emulsion using a Dean Stark moisture trap, also have been employed to remove the liquid phase from the droplet and flocculate the sol.

The second technique induces gelation internally by changing the pH inside the sol droplet. This reduces electrostatic attraction between colloidal species and/or initiates a chemical reaction inside the droplet. For example, urea (H_2NCONH_2) added to a stable sol decomposes and releases ammonia in the aqueous phase during heating of the emulsion. For droplets smaller than 50 μm, amines can be added that diffuse into the droplet and effect an internal pH change. Organic primary amines (R_3CNH_2) added to the sol remain in the organic phase while extracting protons and anions from the droplet, whereas bubbling ammonia diffuses into the droplet to raise the pH. The organic amine route is preferable to the use of bubbling ammonia or dehydration of sol droplets because it minimizes anion contamination in the powder. This is particularly important when the anions present are difficult to eliminate during calcination. In addition, the amine route requires less solvent than the dehydration route, with a 10:1 ratio of gelation solvent to sol for the primary amine route, compared with a 200:1 ratio for dehydration.

"Gel sphere" particles can also be prepared by delivering metal organic monomers to the droplet. First, a hydrochloric acid solution is emulsified in an oil phase (mineral oil and heptane). Silicon tetraethoxide-in-heptane solution is then added to the emulsion to facilitate transport of alkoxide monomers to the acidified droplets where they undergo hydrolysis and polymerization to effect a sol–gel transition within the droplet. This method produces silica gel spheres in the submicron-to-micron range.

Emulsions of nonpolar nonaqueous phases in polar nonaqueous phases can be used to make ceramic powders. Droplets of nonpolar metal alkoxides (or alkoxides dissolved in hexane) can be dispersed and diffusively hydrolyzed and polymerized by the water-soluble polar continuous phase to form spherical powder. Continuous phases such as acetonitrile or propylene carbonate have been employed successfully for insoluble alkoxides such as titanium tetra-*n*-butoxide or aluminum tri-*sec*-butoxide. Systems using propylene carbonate as a continuous phase do not require a surfactant to stabilize the emulsion, whereas those using acetonitrile do. Single- and multicomponent hydrous oxides yield spherical submicron atomically homogeneous oxide powders such as titania, aluminum titanate, or alumina–zirconia. Unlike other emulsion powders synthesized with inorganic salts, these powders are free of anions such as nitrates and chlorides.

3.2 *Microemulsions*

Unlike emulsions, which are kinetically stable, microemulsions are thermodynamically stable. In addition, the dispersed phase of a microemulsion is on a size scale of less than 0.15 μm and therefore is transparent, while emulsions, having a droplet size greater than 1 μm, appear turbid. The reduction in size provides a means of forming finer particles than are possible with emulsions. Both nonoxides and oxides have been prepared from microemulsions. Water-in-oil microemulsions utilizing lyotropic liquid crystals or normal or inverse micelles are usually the systems of interest.

Ultrafine (~4 nm) narrow size distribution magnetite (Fe_3O_4) has been prepared using inverse micelles. The first step is to mix the following two microemulsions: (i) $FeCl_3$ aqueous droplets stabilized with the bisoctyl ester of sodium sulfosuccinic acid in a continuous phase of cyclohexane; and (ii) ammonium hydroxide droplets (microemulsified similarly). Upon the addition of aqueous $FeCl_2$ to the vigorously stirred microemulsion, Fe_2O_3 is formed in the aqueous droplet phase. Precipitation is effected both by diffusion of $FeCl_2$ species and by communication between microemulsion droplets, a process in which two droplets coalesce, exchange contents, and separate, reforming two droplets each with a new composition. Communication allows chemical species (e.g., NH_4OH, $FeCl_2$, and $FeCl_3$) to be exchanged, activating chemical and precipitation reactions in a manner analogous to a titration process. However, not all microemulsion systems rely on communication mechanisms to initiate precipitation. Like emulsions, microemulsion systems can rely on diffusion of reactants through the continuous phase to make a variety of materials which include oxides (SiO_2), carbonates ($CaCO_3$), and borides (Ni_2B).

Reverse microemulsions can also be prepared using supercritical fluids. Submicron aluminum hydroxide particles have been formed by introducing ammonia into a microemulsion of supercritical aqueous aluminum nitrate in propane stabilized by sodium bis(2-ethylhexyl)sulfosuccinate at 100 °C and ~ 20 MPa.

Research has been directed to the preparation of ceramic powders using lyotropic liquid crystals. Monodisperse basic copper sulfate ($[Cu_4(OH)_6]SO_4$) has been prepared from a hexagonal liquid crystal formed by using a mixture of sorbitan monooleate with 20 ethylene oxide groups and aqueous copper sulfate ($CuSO_4$) solution. In this system, precipitation occurs in the interstitial zones between the columns of surfactant molecules arranged in hexagonal arrays. This compartmentalization approach (templating) has also been used to mimic the way nature precipitates ceramic materials such as bone or teeth. For instance, aluminum oxide in the nanometer size range (10–30 nm) has been prepared using single-compartment unilamellar vesicles. A unilamellar vesicle is a hollow sphere with an internal diameter in the range of tens of nanometers. The thin walls facilitate small anion transport but minimize cation diffusion. Vesicles containing aluminum ions are formed by mixing egg yolk phosphatidylcholine ultrasonically with an aqueous solution of aluminum nitrate. After removing all aluminum ions outside the vesicle wall by ion exchange, an aqueous sodium hydroxide solution is mixed with the vesicle suspension. Hydroxyl ions preferentially permeate the vesicle wall to effect a precipitation reaction. A porous synthetic polymer (e.g., poly(dimethylsiloxane) or poly(methylmethacrylate)) can also be used as an alternative to self-assembling surfactants as a template for controlling particle size and morphology. Using a two-stage infiltration process, a hydrolytically sensitive precursor such as a metal alkoxide is introduced into the polymer followed by the addition of a hydrolysis solution. Submicron powders of silica, titania, and barium titanate of sizes with anisotropic and equiaxed morphologies have been produced.

3.3 Aerosols

Aerosols or sprays are liquids dispersed in the gas phase. As with micelle-based systems, particle size control is directly related to the droplet size, which can be controlled by the selection of both the precursor system and the method of aerosol generation.

Submicron aerosols can be generated by vapor condensation. This approach has been used primarily in low-temperature reactions in which particles form within the precursor droplets as they react with the species in the vapor phase. For instance, narrow size distribution 0.2 μm powders can be precipitated from monodisperse 0.4 μm aerosols of alkoxides ($Ti(OR)_4$, Al(O-*sec*-Bu)$_3$) or metal tetrachlorides ($SiCl_4$, $TiCl_4$) by seeding the vapor phase with silver chloride and subsequently hydrolyzing the droplets with water vapor in one or more successive stages at temperatures greater than 0 °C. The hydrolysis is completed by elevating the temperature of the aerosol to 100 °C or more.

Multicomponent systems such as titania–silica powders have been prepared by diffusing $SiCl_4$ vapor into monodisperse $Ti(OEt)_4$ droplets and subsequently hydrolyzing the multicomponent droplet. However, silver chloride contamination and chlorine contamination from the precursor may present problems for ceramics processing applications. Purity problems can be avoided by using a two-stage reaction. First, alkoxide vapor is hydrolytically decomposed to form hydrous oxide seed particles to nucleate an aerosol. This aerosol subsequently can be hydrolyzed to form spherical hydrous oxide particles.

Using mechanical rather than vapor condensation methods for generating droplets, sols have been sprayed into immiscible liquids to form "gel microspheres" by a subsequent dehydration or pH change step as described for emulsions. This technique usually produces particles with sizes in the order of tens of microns. Mechanically generated aerosols have also been used with other types of precipitation chemistries. For instance, an ultrasonic spray atomizer has been used to react aqueous zirconium nitrate with ammonia gas to precipitate a hydrous oxide. Electrostatic aerosols of aqueous solutions of zirconium acetate and polyethylene oxide yielded 0.1–5 μm diameter acetate particles, which were subsequently calcined at 1100 °C to form zirconia. In this system, varying concentrations of polyethylene oxide (< 0.06 wt.%) served as a means to control the particle size. Salts can be spray-dried as well; however, significantly larger particle sizes result. Iron oxides spray dried as various salts typically have particle sizes of 3–20 μm.

Because of their unique solvating properties, supercritical fluids have been used to generate aerosols. With this method, a solute soluble in a supercritical fluid produces supersaturated species by an abrupt loss of solvency as the fluid expands into a low-pressure environment. Selection of the proper pressure and temperature of the pre-expansion and postexpansion conditions produces either one-phase (vapor) or two-phase (liquid–vapor) systems.

Two-phase aqueous systems have been used to precipitate silica, germanium oxide, and uniform mixtures of the two components, while nonaqueous systems have been used to prepare zirconium oxynitrate and polycarbosilane, useful precursors for zirconium oxide and silicon carbide, respectively.

Ceramic powders have been prepared by freeze drying, in which multicomponent liquid solutions are sprayed into a cold liquid to freeze the droplets which are then separated from the liquid and subsequently sublimed. Spherical alumina, magnesium oxide, and spinel powders with submillimeter particle sizes have been prepared from calcined sulfate salts by pneumatically spraying aluminum sulfate solutions into hexane solutions chilled with a dry ice–acetone bath. Acidic barium–yttrium–copper formate solutions have been ultrasonically sprayed into liquid nitrogen ($-196\,^{\circ}C$) to prepare 10–90 μm spheres that calcine as $Ba_2YCu_3O_7$. The results of these two examples demonstrate the importance of the method of aerosol generation in controlling the particle size. Other variables affecting size and the related chemical and physical particle homogeneity, such as spray height distance above the bath and cooling bath temperature, have been examined. In general, it is desirable to spray at a distance that minimizes droplet evaporation and at a temperature that minimizes rejection of the solute as the ice front propagates towards the center of the droplet. In a homogeneous powder, the freezing process imparts a porous, high surface area microstructure to the droplet.

Powders can also be precipitated in spray droplets by solvent replacement. This technique involves the addition of an aerosol of an aqueous solution of metal salts to a nonsolvent that solubilizes water but not the dissolved salts. Various factors, such as the choice of aerosol generator and nonsolvent, control the size distribution and morphology of yttria-stabilized zirconia powder derived from sulfate salts. For instance, acetone produces fused masses, whereas with isopropanol spheres are formed almost exclusively. In ethanol, dropwise addition with a pipette yields a precipitate with a diameter (214 nm) five times that produced by atomization (42 nm).

In spray pyrolysis liquid droplets are sprayed into a reactor that has a low-temperature chamber for precipitating salts and a high-temperature chamber for subsequent pyrolysis. Powders range from micrometer-sized shell-like aggregates to narrow size distribution submicron solid spheres. A broad range of oxide, nonoxide, and composite powders (e.g., WC/Co) are prepared using this approach.

The problems inherent in spray pyrolysis are magnified due to the fact that reactor and precursor systems must incorporate the ideal attributes of both spray-drying and calcination processes. During drying, the kinetics of both solvent evaporation and salt precipitation propagating inwards from the surface of the droplet should be optimized while maintaining the spherical morphology. During pyrolysis, attention must be paid to the large amounts of volatile products that are associated with rate-limiting processes such as heat transfer, diffusion through a product layer, and crust formation. In addition, in each case it must be determined whether or not spray pyrolysis is a liquid–gas phase reaction or a vapor-phase process using a liquid precursor (see *Vapor Synthesis*#). If the latter process is occurring, particle size will not be controlled by droplet size.

4. Summary

Ideal physicochemical characteristics of a ceramic powder include submicron particle size, controlled particle size distribution, uniform and equiaxed shape, minimum degree of agglomeration, high degree of chemical purity, controlled chemical and phase homogeneity, low intraparticle porosity, minimum weight loss upon heating, and controlled surface chemistry. Present technology offers the investigator many alternative approaches for preparing powders with ideal or near-ideal characteristics. There are many technologies to choose from, which enables engineers to make virtually any ceramic composition with the desired physical and chemical characteristics. This article has presented three primary categories of technologies: (i) chemistry in the bulk liquid phase; (ii) chemistry between heterogeneous phases; and (iii) chemistry in a droplet. These technologies center around the site where the chemical reaction to form particles takes place.

Acknowledgments

The author would like to acknowledge the support of the Center for Ceramic Research, Office of Naval Research, and New Jersey State Commission on Science and Technology in preparing this manuscript.

Bibliography

Brinker C J, Clark D E, Ulrich D R (eds.) 1984 *Materials Research Society Symposia Proceedings,* Vol. 32. Materials Research Society, Pittsburgh, PA

Brinker C J, Clark D E, Ulrich D R (eds.) 1986 *Materials Research Society Symposia Proceedings,* Vol. 73. Materials Research Society, Pittsburgh, PA

Hench L L, Ulrich D R (eds.) 1984 *Ultrastructure Processing of Ceramics, Glasses and Composites.* Wiley, New York

Hench L L, Ulrich D R (eds.) 1986 *Science of Ceramic Chemical Processing.* Wiley, New York

Mackenzie J D, Ulrich D R (eds.) 1988 *Ultrastructure Processing of Advanced Ceramics.* Wiley, New York

Messing G L, Mazdiyasni K S, McCauley J W, Haber R A (eds.) 1987 *Advances in Ceramics: Ceramic Powder Science,* Vol. 21. American Ceramic Society, Westerville, OH

Onoda Jr G Y., Hench L L (eds) 1978 *Ceramic Processing Before Firing*. Wiley, New York
Pugh R J, Bergström L (eds.) 1994 *Surface and Colloid Chemistry in Advanced Ceramics Processing*. Dekker, New York

R. E. Riman
Rutgers University, Piscataway, New Jersey, USA

Processing Flexible Polymers to High Performance Fibers

In the 1970s, significant progress was made in exploiting the intrinsic properties of the macromolecular chain, especially in the field of 1D objects such as fibers. Two major routes can be discerned which are completely different in respect to the starting (base) materials, namely rigid as opposed to flexible macromolecules (Lemstra *et al.* 2000). The prime examples of rigid chain polymers are the aromatic polyamides (aramids), notably poly(*p*-phenylene terephthalamide) (PPTA), currently produced under the trade names Kevlar (Du Pont) and Twaron (Akzo Nobel). More recent developments include the PBO fiber from Toyobo, Zylon, and the experimental fiber M-5 from Akzo Nobel based on poly-diimidazo pyridynylene dihydroxy phenylene (Lammers *et al.* 1998). The *primus inter pares* of a high-performance fiber based on flexible macromolecules is polyethylene (PE) which is currently produced by DSM (Dyneema) and its licensee Allied Signal (Spectra). The main difference between rigid and flexible chains is the necessity to force chain extension in the case of flexible polymer molecules in order to exploit the intrinsic possibilities of the chain concerning ultimate mechanical properties, while for rigid rod polymers chain-extension has been built in by synthesis. Figure 1 shows the specific, based on performance by weight, strength versus specific modulus of the various fibers.

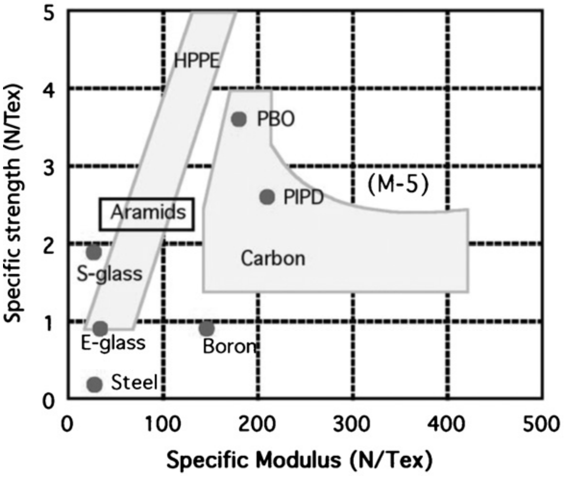


Figure 1
Specific strength versus specific modulus of various fibers. HPPE = high-performance PE fibers.

1. Ultimate Properties

1.1 Tensile Modulus

The ultimate tensile moduli of PE and other polymer systems, can be calculated or obtained experimentally from x ray diffraction measurements on oriented fibers during mechanical loading (Nakamae and Nishino 1992). All the literature data show that the Young's modulus E of PE in the chain direction is extremely high, $\geqslant 200$ GPa, due to the small cross-sectional area of the chain, the absence of side groups, and the planar zigzag chain conformation in the orthorhombic crystal lattice. PE is not unique and other linear polymers with similar crystal structures possess similar theoretical moduli, e.g., poly(vinyl alcohol). In contrast, isotactic polypropylene has a lower ultimate modulus (60 GPa) in the chain direction due to its 3_1 helix conformation of the chain in the crystal lattice.

1.2 Ultimate Tensile Strength

In the past, a variety of studies have been devoted to the theoretical tensile strength of oriented and chain extended structures, i.e., the breaking of chains upon loading (He 1986). The theoretical tensile strength of a single, extended polymer chain can be calculated directly from the C–C bond energy. These calculations show that the theoretical tensile strengths are extremely high, in the order of 20–60 GPa (at 0 K). This requires a defect-free, chain-extended structure and *infinite* polymer chains. In practice, however, we are dealing with *finite* chains and a completely different situation is encountered as will be addressed in sect.2.3.

1.3 Infinite Versus Finite Chains

Upon loading an array of perfectly aligned and extended finite polymer chains, the stress transfer in the system occurs via secondary, intermolecular bonds. Chain overlap is needed in order to be able to transfer the load through the system. Qualitatively, it could be envisaged that the bonds in the main chains are only activated when the sum of the small secondary interactions, $\Sigma\varepsilon_i$, approaches the bond energy, E_i, in the main chain. Intuitively, one expects that in order to obtain high-strength structures in the case of PE, a high molar mass is needed, in combination with a high degree of chain-extension, to build up sufficient intermolecular interactions along the chains. Termonia and Smith (1987) used a kinetic model to simulate the fracture behavior of an array of aligned and extended *finite* polymer chains. It was found that the molecular weight (or number of chain ends) has a profound influence both on the fracture mechanism and on the theoretical tensile strength of these hypothetical structures. In order to obtain a strength level of 5 GPa, a molar mass of $>1\times10^5$ Dalton is needed for PE whereas 5×10^3 Dalton is sufficient for PPTA. The conclusion is that polymers possessing strong secondary bonds require a smaller overlap length to obtain a high tenacity.

2. *Melt-spinning/Drawing*

Chain orientation and extension is generated in melt- and solution-spun fibers by: (i) applying a draw-down to the fibers during or immediately after spinning (in the molten state or super-cooled melt) and (ii) drawing of fibers at temperatures close to, but below, the melting or dissolution temperature. Drawing in the (semi-)solid state, i.e., below the melting and/or dissolution temperature is usually much more effective. Relaxation processes are restricted since the chains are trapped into crystals which act as physical network junctions. In the 1970s, Capaccio and Ward (1974) started systematic studies concerning the drawability of linear PEs in the solid state and they developed a technological route. By optimizing the polymer composition and process conditions, PE fibers could be produced possessing tensile (Young) moduli up to 75 GPa and a strength level up to 1.5 GPa. A well-known observation made by Capaccio and Ward (1974) based on numerous *isothermal* drawing experiments, is that with increasing molar mass the maximum draw ratio decreases towards a limiting value of 4–5 at MW values over 1×10^6 Dalton. A limited drawability in the solid state is not unique for PEs. Many other polymers demonstrate a limited drawability, for example polyamides, which are often referred to as the natural draw ratio.

To understand the drawing behavior of PEs in the solid-state, one automatically focuses on the role of crystallites, viz the folded-chain crystals which are organized in more or less well-developed spherulites in melt-crystallized samples. There is, however, no direct correlation between crystal size/crystallinity, or morphology in general, versus the maximum draw ratio. Depending on the crystallization conditions, the drawability can change dramatically for one specific sample. For example, slow cooling from the melt can promote the drawability in the solid state Capaccio and Ward (1974) but also can cause identical PE samples to become brittle in the case of very slow cooling from the melt. In conclusion, melt-spinning/drawing is limited with respect to the molar mass of the PEs. With increasing molar mass, both the spinnability (a strong increase of melt viscosity) and the drawability in the solid-state decrease, which sets an upper limit to melt-spinning of PEs of typically 500 kg mol^{-1}.

3. *Solution-processing and Solution(gel)-spinning*

An obvious route to increase the spinnability of high molar mass PEs is to use solvents to lower the viscosity. In retrospect, the early attempts in the 1960s to employ solution-spinning of high molar mass PEs (Juergenleit 1956), to produce high-strength/high-modulus fibers failed since no proper understanding was available at that time concerning optimized drawing of solution-spun fibers. At the end of the 1970s, solution (gel)-spinning of UHMW-PE was discovered at DSM (DSM/Stamicarbon 1979, Smith and Lemstra 1979). In the solution (gel)-spinning technique, semi-dilute solutions are employed during spinning but the elongation of chains is performed by drawing in the semi-solid state, i.e., below the melting (dissolution) temperature. A solution of UHMW-PE with a low polymer concentration, typically of 1–2%, was spun into water.

Upon cooling a gel-like filament is obtained consisting of a physical network, obtained by thermoreversible gelation, containing a large amount of solvent. The as-spun/quenched filaments are mechanically sufficiently strong (gel-fibers) to be transported into an oven in which drawing is performed. The ultra-drawability of these gel-fibers seems not too surprising in view of the large amount of solvent which could act as a plasticizer during the draw. The remarkable feature, however, is that ultra-drawing is still possible after complete removal of the solvent prior to the drawing process. The solvent is necessary to facilitate processing of the rather intractable polymer UHMW-PE and induces a favorable structure/morphology for ultra-drawing but the solvent is not essential during the drawing process.

4. *The Entanglement Concept*

A very simple model for the enhanced drawability of solution-spun/cast UHMW-PE was put forward by Smith *et al.* (1982) based on a network approach, ignoring completely the morphology and crystal structure. In this model it is assumed that entanglements are trapped in PE upon crystallization and act as semi-permanent cross-links in a physical network upon solid state drawing. The crystals as such do not resist deformation in the temperature domain, close to but below the melting temperature, which is employed for (ultra-) drawing. The maximum draw ratio $\lambda_{\max}$ scales with the ratio between a fully stretched strand, scaling with M_e, between two entanglement points, and the average unperturbed end-to-end distance which scales with the square root of M_e, hence

λ_{max} scales with $M_e^{0.5}$. Upon dissolution, the entanglement density decreases and the molar mass between entanglement loci increases, in good solvents about proportional to the inverse of the volume fraction of polymer in solution: M_e (solid) = M_e (melt)/φ. The maximum draw ratio of solution-cast/spun PEs is consequently given by: $\lambda_{max} \propto \varphi^{-0.5}$. For example, as-spun fibers from a 1% solution have an increased drawability with a factor of 10.

The entanglement model is remarkably versatile and can explain various phenomena such as the limited drawability of melt-crystallized UHMW-PE, λ_{max} 4–5, the dependence of the maximum draw ratio of solution-cast/spun UHMW-PE on the initial polymer concentration in solution, $\lambda_{max} \propto \varphi^{-0.5}$, and the increase of the drawability in the solid state upon isothermal crystallization at low supercoolings of UHMW-PE solutions, or from the melt in general, due to the fact that the chains are reeled in, viz pulled out their entanglement network. It is very important to note that the proposed ‘‘entanglement network’’ model is not universally valid. It can be applied to apolar polymers such as PEs and polypropylenes, but not to polymers possessing relatively strong secondary interactions, such as hydrogen bonds. In the case of polyamides, the folded chain crystal resist deformation and unfolding (Postma *et al.* 1990).

5. *Solvent-free Processing of UHMW-PE; Nascent Reactor Powders*

For solution(gel)-spinning of UHMW-PE a lot of solvent is needed for disentangling which has to be fully recovered in view of increased environmental legislation. Consequently, the challenge is to find ‘‘solvent-free’’ routes starting from disentangled precursors. The rationale behind this approach is that once disentangled UHMW-PE structures are obtained via some route, subsequent melt-processing should become feasible. One would at least expect a time-temperature window in which disentangled UHMW-PE should possess a lower initial melt-viscosity in comparison with a standard equilibrium melt. This is in accordance with the long relaxation (reptation) times (1×10^4 seconds per 200 °C) of UHMW-PE, viz complete re-entangling is a time consuming process. To prepare (completely) disentangled UHMW-PE structures is feasible and rather straightforward using modern catalysts. Polymerization conditions are known, viz low temperature and rather low catalyst activity, which promote the formation of folded-chain crystals directly on the surface of the (supported) catalyst (Lemstra *et al.* 2000). In the limit of a low concentration of active sites on the catalyst (surface), one could expect that the individual growing chain would form its own *monomolecular* crystal. Despite all efforts made to prepare specific disentangled UHMW-PE precursors for subsequent melt-spinning, the ultimate conclusion at this point in time is that processing disentangled UHMW-PE with the aim to benefit from an initial lower melt-viscosity and to preserve the disentangled state to some extent during processing and prior to drawing, is not feasible at all. The salient feature is that disentangled UHMW-PE, either obtained by collecting precipitated single crystals or via specific low-temperature polymerizations, shows (Engelen and Lemstra 1991) the same high melt viscosity (in shear) upon heating above the melting temperature as standard ‘‘equilibrium’’ UHMW-PE melts. No memory effect from any previous polymerization/crystallization history can be depicted, and moreover, upon re-crystallization from the melt, the favorable drawing characteristics of disentangled UHMW-PE are lost completely and the drawing behavior is indistinguishable from a standard melt-crystallized UHMW-PE sample.

This intriguing problem of no memory effect and immediate loss of drawability upon re-crystallization has been addressed by Sadler and Barham (1990) following the fate of a PE chain using neutron scattering upon melting. They showed that the radius of gyration, R_g, which is rather low in solution-crystallized samples due to chain folding, immediately jumps to its ‘‘equilibrium’’ random coil dimensions upon surpassing the melting temperature (*chain explosion*). The experiments concerning the fast decay in drawability upon re-crystallization, related to the phenomenon of ‘‘*coil explosion*,’’ point out that rather local changes, involving chain segments rather than complete chain motion/diffusion (reptation), can dominate the drawing behavior and these results prompted us to modify the ‘‘entanglement model’’ to some extent (Lemstra *et al.* 2000). Upon crystallization from solution, the molecules fold usually along the {110} plane and the stems of a test chain (heavy dots) are shown in Fig. 2 without indicating the folds. For the sake of simplicity, we assume that adjacent re-entry occurs during crystallization and that the chain is located within one crystal plane. Shearing and unfolding in the direction perpendicular to the chain and along the {110} plane is rather easy in view of the low shear moduli. Upon melting these crystals the chains will immediately adopt a random coil conformation and stems of different molecules will interpenetrate in the ‘‘coil explosion’’ process. Upon re-crystallization, the stems of the test chain become crystallized in a more random order within the crystal and shearing (slip) is more difficult since the chains cannot cross mutually during deformation (Fig. 2).

The instantaneous loss in drawability upon melting and re-crystallization is due to re-arrangement and intermixing of stems involving only local chain motions rather than movement of the complete chains chain as proposed for self-diffusion in polymer melts. The driving force for the mixing of stems and coil-explosion upon melting is the gain in entropy (the

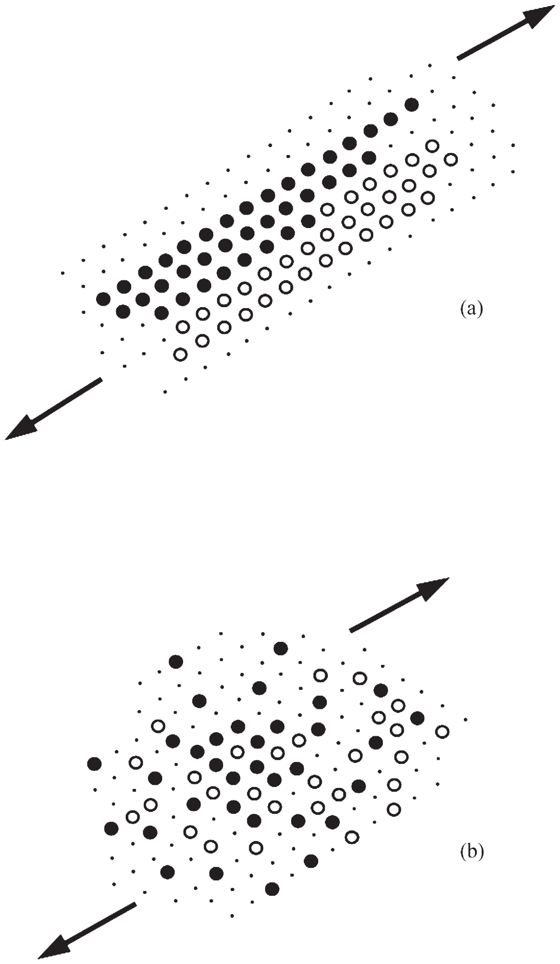


Figure 2
Stem arrangement in PE crystals, adjacent (a) versus re-entry (b).

entropy of fusion which is absent in the diffusion/reptation processes in quiescent polymer melts.)

6. *Solid-state Processing*

From sect.5, it may be concluded that all attempts to prepare disentangled UHMW-PE structures by specific polymerizations are in vain since no advantage could be obtained in the subsequent melt-processing. This conclusion is, in fact, not true. The only lesson to be learned is that the disentangled UHMW-PE precursors should never be heated above the melting temperature. Below the melting temperature, the disentangled UHMW-PE reactor powders, the so-called nascent or virgin powders, are remarkably ductile and can be processed via calandering or hot-compacting and subsequently drawn into tapes or fibrillated structures. The draw ratio of well-prepared nascent UHMW-PE powders is similar to solution-crystallized UHMW-PE samples. Processing of UHMW-PE reactor powders has been partly successful for making oriented tapes by sintering/compacting between rollers and subsequent drawing. Kanamoto *et al.* (1987) developed a two-stage drawing process for reactor powders and they obtained Young's moduli of over 100 GPa. Nippon Oil Company developed and patented several solid-state processing routes for making strong UHMW-PE tapes. The only problem with solid-state processing is the handling and limited flow of powders. It is not possible to produce fine denier filaments. To enhance the flowability processing via the hexagonal phase might be a future option.

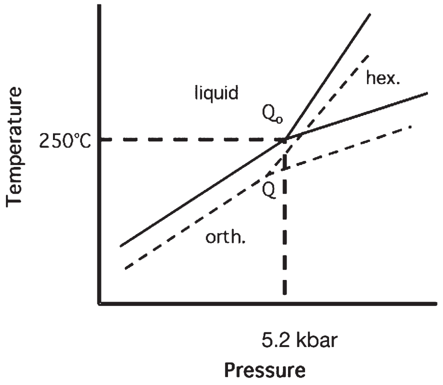


Figure 3
P–T diagram of PE.

7. *Processing via the Hexagonal Phase*

The hexagonal phase of (isotropic) PEs is usually observed at elevated pressure and temperature. The (equilibrium) triple point *Q* is located at 5.2 kbar and 250 °C (see Fig. 3). It was shown by Rastogi *et al.* (1998b) that this triple point is dependent on the crystal size/dimensions and can be lowered to approximately 1 kbar in the case of nascent UHMW-PE reactor powders which possess very small dimensions (<10 nm). This opens the possibility for processing via the (mobile) hexagonal phase. In the hexagonal phase, the chains are mobile but only in the chain direction via chain-sliding, $\langle c \rangle$, or diffusion. Upon annealing in the hexagonal phase ultimately complete chain extension will be obtained and extended-chain crystals will be obtained. Since there is no motion perpendicular to the chains, the drawability is not lost based on our revised model for drawing (see Fig. 2). In fact, the drawability can even be enhanced in the hexagonal phase as shown by Maxwell *et al.* (1996)

since additional "disentangling" will occur as a consequence of extended-chain crystal formation. To date no processing routes via the hexagonal phase for making PE fibers are known. However, processing via the hexagonal phase for making hip joints based on UHMW-PE have been discussed in the literature (Rastogi *et al.* 1998b).

8. Drawing Behavior of Other Flexible Polymer Systems

The success of melt-spinning/drawing in general, and of solution(gel)-spinning in the particular case of ultra-high molecular weight PE, to produce oriented PE structures possessing high tensile moduli, stimulated research activities concerning the drawing behavior of other linear polymers, notably polypropylene, polyoxymethylene, and the aliphatic polyamides, nylon 6 and nylon 66. The high melting temperatures, 225 °C (nylon 6) to 310 °C (nylon 46) compared with PE (~150 °C), and the presence of hydrogen bonds, which could reduce the creep make the polyamides attractive candidates. Extensive research effort to produce high modulus and strength fibers based on aliphatic polyamides have failed, however, despite major efforts in industry to perform solution(gel)-spinning based on high molar mass polyamides. The reason for this failure was demonstrated by Postma *et al.* (1990) who showed that the hydrogen bonds in lamellar, solution(gel)-crystallized polyamides, are essentially static up to the melting temperature and act as barriers, thus prohibiting draw. Solution(gel)-spinning promotes the formation of rather well-developed polyamide (single) crystals and, consequently, unfolding into extended-chain structures becomes more difficult.

9. Properties of PE Fibers

9.1 Tensile Strength (1D)

Currently, UHMW-PE fibers possessing tensile strengths of 3–4 GPa are produced commercially, e.g., Dyneema by DSM and Spectra by its licensee Allied Signal (see Fig. 1). Due to its low density, the specific values for the stiffness and strength of PE fibers are currently superior, at least at *ambient* temperatures.

9.2 Melting Temperature

The melting temperature is given by $T_m = \Delta H/\Delta S$ and the chains in the fiber can not melt when the fiber is constrained, because the chains cannot adopt a random coil conformation. Consequently, ΔS is low and the melting temperature increases. In the case of UHMW-PE fibers, this enhanced melting temperature by external constraints is not possible due to the onset of the hexagonal phase at 155 °C. Above this temperature, the orthorhombic crystals in the fiber transform into the hexagonal structure. In the hexagonal phase the chain mobility is high and the fiber cannot sustain any load and fails (Lemstra *et al.* 2000).

9.3 3D Properties

PE fibers are highly anisotropic structures. The favorable characteristics in the fiber (chain) direction in terms of stiffness and strength are not matched at all in off-axis properties. The low transverse and shear moduli of PE are due to the absence of specific interactions perpendicular to the chain (only weak van der Waals bonding). The absence of strong interactions between adjacent chains causes the strength of the fibers to also be highly anisotropic. This is a major drawback in structural composite applications. It was shown by Peijs (1994) that the low interlaminar shear strength (ILLS) obtained from a three point bending test of unidirectional composites containing UHMW-PE fibers (Table 1) is inherently due to the poor shear and compressive properties of PE fibers. Consequently, UHMW-PE fibers cannot be used in structural composite applications.

9.4 Long-term Mechanical Properties: Creep

Soon after the gel-spun fibers were produced, it became apparent that the exceptionally good short-term mechanical properties were not matched by equally good long-term properties. In tensile tests at low strain rates, the UHMW-PE fibers exhibit a drop in Young's modulus and they exhibit a yield stress which is highly dependent on the strain rate. In fact, it is shown that the tensile strength (yield stress) drops by at least an order of magnitude at prolonged time scales. Some representative creep data of UHMW-PE fibers, in this case of a Dyneema fiber (SK66) possessing a tensile modulus of ~100 GPa are shown in Fig. 4. Figure 4 shows that initially the elongation increases proportionally to the logarithm of time.

The creep becomes important at longer loading times. This pronounced, irreversible creep is not too

Table 1
Interlaminar shear strength (ILSS) of unidirectional composites; 50/50 fiber/epoxy composites.

Fiber	ILSS (MPa)
UHMW-PE (untreated)	13
UHMW-PE (corona/plasma treated)	20–30
Aramid	45–70
E-glass	75–95
Carbon	80–120

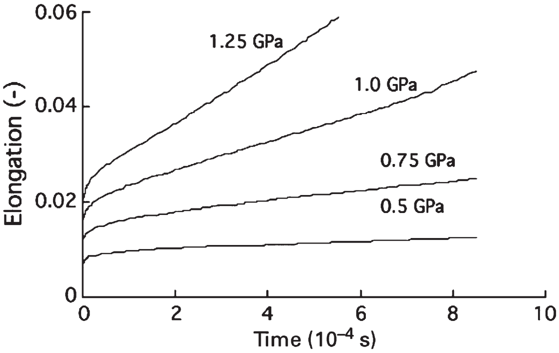


Figure 4
Elongation (creep) of an UHMW-PE fiber at ambient temperature. The Young's modulus (E) of the fiber is $\sim$100 GPa.

Table 2
Major applications of solution(gel)-spun PE fibers.

Ropes and cables	Ballistic protection	Miscellaneous
Towing lines	Bullet proof vests	Sails
Mooring lines	Insert for vests	Motor helmets
Yacht ropes	Helmets	Cut resistant gloves
Long lines	Car armor panels	Radomes
Trawl nets	Spall liners	Dental floss
Fish farms	Ballistic blankets	Speaker cones
Parapent lines	Containment shields	Cryogenic composites

surprising when we take into account that UHMW-PE fibers are made by ultra-drawing and in the drawing process folded-chain crystals are transformed into chain-extended structures. In order to avoid fiber breakage in the drawing operation, the fibers are never drawn to their maximum draw ratio $\lambda_{\max}$ and, since there is no "lock-in" mechanism, the fibers can de drawn further in any subsequent tensile loading experiment, such as a creep test. Consequently, there is no basic difference between the drawing operation and a creep test, especially at elevated temperatures. All known methods for improving the creep resistance such as cross-linking (e.g., electron beam and UV) and chemical modification, require the presence of a certain degree of imperfection in the fiber in order to bring the chains together to form a covalent bond. Creep improvement, therefore, implies a compromise between short-term and long-term mechanical properties (Lemstra *et al.* 2000).

10. Major Applications of PE Fibers

In the previous sections the properties of solution-spun, ultra-drawn UHMW-PE fibers were discussed in terms of their tensile (Young) modulus and tensile strength which show impressive values at short loading times but rather extensive creep at prolonged loading times. These properties are, of course, relevant in a large number of applications such as fishing nets and/or lines, ropes, and composites. Such fibers also possess a number of favorable properties such as a high-impact resistance, high-cutting resistance, low dielectric constant, low dielectric loss, low stretch, high heat conductivity in the fiber direction, and a high sonic modulus. These "secondary" properties are particularly important in applications such as ballistic applications (bulletproof vests), helmets (impact), hybrid composites (impact), gloves (cutting resistance), fishing lines (low stretch), loudspeaker cones (sonic modulus), and radomes/sonar domes (dielectric properties). The properties which motivate the use of highly oriented gel-spun UHMW-PE fibers are high (specific) strength and modulus (all applications), low density, high energy absorption capability (ballistic protection and composites), high sound speed (ballistic protection and speaker cones), flexibility, (ropes and nets), good dielectric properties (radomes), and good chemical resistance (Table 2).

See also: Polymer Fiber Processing: Modeling#; Polymer Fiber Processing: High-performance Fibers*; Processing Rigid Polymers to High Performance Fibers*; Solid-state Extrusion to High Performance Fibers#

Bibliography

Bastiaansen C W M, Meijer H E H, Lemstra P J 1990 Memory effects in polyethylenes–influence of processing and crystallization history. *Polymer* **31** (8), 1435–40
Blades H, White J R 1963 *US Pat.* 3 081 519
Capaccio G, Gibson A G, Ward I M 1979 *Ultra-High Modulus Polymers*. Elsevier Applied Science, London
Capaccio G, Ward I M 1974 *Polymer* **15**, 233
Capaccio G, Ward I M 1975 *Polymer* **16**, 243
DSM/Stamicarbon 1979 *US Pat.* 4 137 394
He T 1986 *Polymer* **27**, 253
Juergenleit W 1956 *US Pat.* 3 048 465
Kanamoto T, Ohama T, Tanaka K, Takeda M, Porter R S 1987 *Polymer* **28**, 1517
Kelly A, Macmillan N H 1986 *Strong Solids*. Clarendon, Oxford, p. 22
Lammers M, Klop E A, Northolt M G, Sikkema D J 1998 Mechanical properties and structural transitions in the new rigid-rod polymer fibre PIPD ('M5') during the manufacturing process. *Polymer* **39** (24), 5999–6005
Lammers M 1998 Ph.D. thesis, PIPD rigid-rod polymer fibres and films; structure formation and mechanical properties. Swiss Federal Institute of Technology (ETH), Zurich

Lemstra P J, Bastiaansen C W M, Rastogi S2000 In Salem D R *Structure Formation in Polymeric Fibers*. Hanser, Munich, Germany 185-223. ISBN 3-446-18203-9

Maxwell A S, Unwin A P, Ward I M 1996 The effect of high pressure annealing on the molecular network in polyethylene. *Polymer* **37** (15), 3293–301

Nakamae E K, Nishino T 1992 *Adv. X-Ray Analysis* **35**, 545

Nishino T, Ohkubo H, Nakamae K 1992 Studies on the temperature-dependence of the elastic-modulus of crystalline regions of polymers. 12. Temperature-dependence of the elastic-modulus of crystalline regions of polyethylene with different microstructures–explanation with the kinked-chain model. *Journal of Macromolecular Science-Physics* **31** (2), 191–214

Peijs A A M 1994 Ph.D. thesis Eindhoven University of Technology, High-performance polyethylene fibres in structural composites. ISBN-90-386-0033

Postema A R, Smith P, English A D 1990 Ultra-drawing of polyamides – the hydrogen-bond barrier. *Polymer Communications* **31** (12), 444–7

Rastogi S Koets P, Lemstra P J 1998a *NL Pat. Applic 98/00093*. Method for processing polyethylene having a molecular weight of at least 400.000 by heating at an elevated pressure

Rastogi S, Kurelec L, Lemstra P J 1998b Chain mobility in polymer systems; on the borderline between solid and melt 2. Isothermal phase reversal and sintering of ultra high molecular weight polyethylene. *Macromolecules* **31**, 5022–31

Sadler D M, Barham P J 1990 Structure of drawn fibers. 1. Neutron-scattering studies of necking in melt-crystallized polyethylene (English). *Polymer* **31** (1), 36–42

Smith P, Lemstra P J 1979 *Br. Pat.* 2 051 661

Smith P, Lemstra P J, Booij H C 1982 *J. Polym. Sci. Phys. Ed.* **20**, 2229

Termonia Y, Smith P 1986 *Polymer* **27**, 1845

Termonia Y, Smith P 1987 *Macromolecules* **21**, 835

Tervoortengelen Y M T, Lemstra P J 1991 Morphology of nascent ultra-high-molecular-weight polyethylene reactor powder–chain-extended versus chain-folded crystals. *Polymer Communications* **32** (11), 343–5

Zwick M 1965 *NL Pat. Applic.* 6 501 248

P. J. Lemstra and S. Rastogi
Dutch Polymer Institute/Eindhoven University of Technology, Eindhoven, The Netherlands

Processing Rigid Polymers to High Performance Fibers

Fiber materials obtained from natural sources have been used through the ages for reinforcement and other structural applications. Well-known examples include the use of straw as reinforcement of clay and plaster, and the use of flax and hemp for cables and ropes. Although such natural fibers are still used extensively, high performance man-made fibers take a steadily increasing share of the market. This is due to the fact that they allow the production of continuous fibers with exceptional mechanical properties, for instance specific tensile strengths that are 5–10 times higher than that of steel. In addition to the purely synthetic materials, also natural sources for man-made fibers are of major importance, e.g., regenerated cellulose. To obtain fiber materials with high performance properties requires processing methods that yield extremely high molecular orientation of the polymer chains along the fiber axis. For the processing of rigid polymers various techniques are possible. With the first method the required high level of molecular orientation is obtained by the processing of rigid polymer chains from liquid-crystalline (LC) nematic solution, using the dry-jet, wet-spinning technique. Alternatively, rigid polymers can be spun from isotropic solutions. In this case a relatively low degree of orientation is induced during the spinning process. The high orientation, which is required for a high tenacity and modulus is imposed in a hot drawing after-treatment. Kevlar (DuPont), Twaron (Acordis), and Zylon (Toyobo) are representatives of the first category, Technora (Teijin), Armos, and Rusar of the second. Because, from an industrial point of view the dry-jet, wet-spinning of nematic solutions is the most important one, this article will concentrate on this process route. Nematic LC spinning solutions are liquids where the polymer chains show spontaneous alignment along a locally defined axis of orientation $\tilde{n}$, the director (see *Liquid Crystalline Polymers: An Introduction*$^{\#}$). Figure 1 shows the molecular arrangement in a nematic polymer solution.

1. History

The first observation of a LC rigid polymer solution was reported by Conmar-Robinson for poly(benzyl-*L*-glutamate) or PBLG in m-cresol, subsequently also p-aramids were found to show such behavior in concentrated sulfuric acid solution. The latter solutions were found to be suitable for the spinning of high performance aramid fibers, presently known as Kevlar (DuPont) and Twaron (Teijin Twaron). Inspired by this success, alternatives to aramid spinning were developed: PBO and PBT fibers, and more

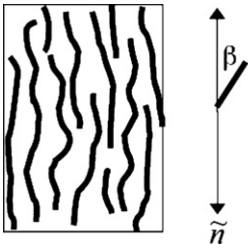


Figure 1
Nematic liquid-crystal structure of a main-chain polymer solution, the angle β indicates the deviation of the molecular chain tangent from the average director $\tilde{n}$.

recently PIPD (or M5) and a new cellulose yarn. The last example refers to a spinning process that makes use of nematic cellulose solutions and is an excellent example of a high performance man-made fiber from natural sources.

2. *Chemical Structures*

The global requirements for polymer chains that are suitable for fiber spinning from nematic solution are rigid chains, a small molecular cross-section, and strong lateral interactions. Some examples of molecular structures that fit this description are shown in Fig. 2, together with the solvent of choice.

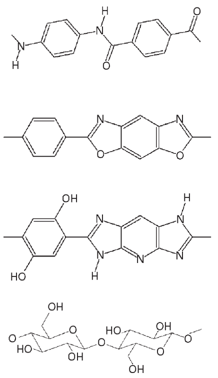
Clearly it is necessary to identify solvents suitable for use in a dry-jet wet-spinning process, where the solvents are preferably cheap and not highly toxic. Contrary to common belief strong acids such as sulfuric acid and phosphoric acid while being corrosive are not particularly toxic or volatile, and, therefore, can be used on an industrial scale without putting the operating personnel at any undue risk. For this reason other known solvents, such as hydrofluoric acid for aramids and trifluoroacetic acid for cellulose-acetate, are not generally considered to be serious options for commercial spinning processes.

3. *Liquid-crystalline Order*

The remarkable mechanical properties of fibers, made from the above polymers, can be understood from the occurrence of LC order in the spinning solution (see Fig. 1). Due to the stiffness of the polymer chains in these materials there is a tendency for them to pack together in a uniform fashion. This so-called LC order manifests itself in the fact that at a given temperature and for sufficient polymer concentration the polymer solutions show distinctly anisotropic properties. The anisotropy relates to many observable physical properties, such as the optical (birefringence) and the rheological properties, where the collective molecular alignment causes a dependence on the orientation of the sample with respect to the external fields. Many LC phases are known, with varying degrees of orientational and positional order. For reasons of simplicity this article will be confined to the simplest case of the nematic phase. This is a fluid phase where the molecules show collective orientation but without the positional order commonly associated with true crystalline materials. The LC structure shown by these polymer solutions does strongly depend on the polymer concentration and the temperature. For a given polymer concentration, at high temperatures the polymer solution will be isotropic, i.e., without orientation, whereas at low temperatures the molecules will show the collective nematic order described above.

At a well-defined temperature the material will show a phase transition from the nematic to the isotropic phase, which is similar in nature to better known phase transitions like the freezing and boiling point. Figure 3 shows this so-called clearing temperature for three polymer solutions.

In addition to this, the degree of molecular orientation in the nematic phase will depend on temperature, at low temperatures the orientational order will be more perfect than at higher temperatures. This is obvious when the thermal motion of the polymer chains is taken into account. The degree of orientation can be expressed as a number, the orientational order parameter, which by definition is zero for an isotropic phase and is one for a material with perfect molecular alignment. It should be clear that the order parameter will increase with increasing polymer concentration and will decrease with increasing temperature. This can be made more quantitative using a variety of theoretical models such as the Maier-Saupe model and the Onsager model.



PPTA: poly (para- (phenylene terephtalamide) Twaron, Kevlar
polymerization solvent NMP/$CaCl_2$
spinning solvent H_2SO_4

PBO: polybenzoxazole
Zylon
polymerization in and spinning from phosphoric acid

PIPD or M5: poly{2, 6-diimidazo[4,5-b: 4′,5′-e]pyridinylene-1,4-(2, 5-dihydroxy)phenylene}
polymerizattion and spinning from phosphoric acid

Cellulose
spinning from phosphoric acid

Figure 2
Structures of some rigid main chain LC polymers.

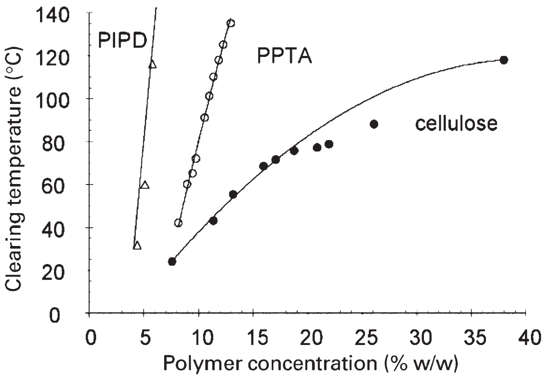


Figure 3
Clearing temperature from the nematic to the isotropic phase of PPTA in sulfuric acid, and PIPD and cellulose in phosphoric acid.

4. Spinning of Liquid-crystalline Polymer Solutions

Poly(p-phenylene terephthalamide) (PPTA), poly-*p*-phenylenebenzobisoxazole (PBO), poly{2,6-diimidazo[4,5-b:4′,5′-e]pyridinylene-1,4-(2,5-dihydroxy)phenylene} (PIPD), and cellulose are all spun from LC (nematic) solutions. After being filtered and heated, the LC solutions are extruded through gold/platinum spinnerets containing approximately 1000 capillaries with a diameter of 50–200 µm. Figure 4 gives a schematic view of the dry-jet wet-spinning processes.

Rheology of the LC phase plays an important role in processing and structure formation. Shear flow occurs during the melting of the solution, in the pipes of the machine, and in the capillaries of the spinneret. A well-known phenomenon, which was first observed by Kwolek (DuPont) on solutions of aromatic polyamides, is the fact that at low shear rates the shear viscosity exhibits a maximum as a function of the polymer concentration. This maximum in the viscosity versus concentration corresponds to the onset of the nematic phase as is also explained by the Doi theory. When the shear rate increases, the maximum becomes smaller and can even disappear. This is due to flow-induced orientation of the isotropic phase, which diminishes the difference in flow characteristics between the isotropic and the nematic phase. Main-chain LC phase solutions are strongly shear-thinning, especially in the region of the phase transition. Elongational flow fields occur in the entrance zone of the spinneret capillaries and in the air gap, where a draw ratio of approximately 10 is imposed. The elongational viscosity is much less dependent on the deformation rate than is the case for the shear viscosity. For example, with cellulose solutions in phosphoric acid the velocity profile in the air gap has an exponential shape, indicating a constant elongational viscosity. PPTA solutions show similar behavior. The elongational viscosity of PPTA solutions is in the order of 1 kPa, while for cellulose solutions in phosphoric acid it is about 20 kPa.

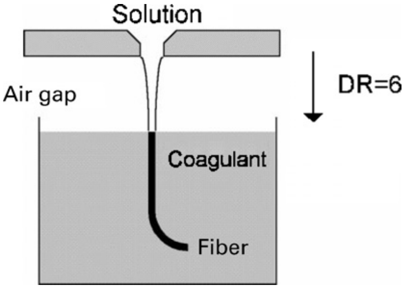


Figure 4
Schematic view of the dry-jet wet-spinning process, the solvent is removed in the coagulation bath containing the coagulant (e.g., H_2O or dilute sulfuric acid for PPTA-spinning).

For some polymers, e.g., PBO, it can be useful to cool the filaments in the air gap by blowing air onto them. This method is also used in melt spinning of cellulose fibers from NMMO solutions (Tencel, Acordis and Lyocel, Lenzing) and can reduce the occurrence of draw resonance.

Compared to solutions or melts of flexible polymers, the die swell for LC solutions of rigid polymers is rather small. The birefringence in the die swell area is very low, but once the velocity in the air gap passes the extrusion velocity the level of birefringence is close to saturation. Additional further drawing slightly improves the orientation of the nematic phase according to the affine deformation model. In this model the orientation of the director field is calculated by assuming that the strain in the liquid-crystal solution is uniform and equal to the externally applied strain. This allows the calculation of the director orientation in terms of a ‘‘director order parameter’’ that is a function of the externally applied strain only. This director order parameter can be used together with the local orientational order parameter, of the molecules with respect to the director, to calculate the mechanical properties of spun fibers as a function of temperature, concentration, and draw-ratio.

The high level of orientation, which is created during the dry-jet spinning, should be maintained in subsequent washing and drying steps. This is achieved by keeping the yarn under tension in these process steps.

5. Structure of Main Chain Liquid-crystalline Phase Spun Fibers

In general the rigid rod polymeric chains of the LC phase fibers bring about a highly crystalline and oriented structure, which is borne out by the results of x ray diffraction, electron diffraction, and in particular by the lattice imaging technique. The chains in cellulose II, PPTA, and PIPD are packed in a 3D order extending in transverse direction over 5–8 nm and in longitudinal direction over more than 40 nm. In the PBO structure the 3D order is disturbed by random translations parallel to the chain axis. In cellulose II and PIPD the chains are laterally bonded by hydrogen bonds in two directions, whereas the crystalline structure of PPTA the hydrogen bonding between the chains extends in a single direction. Different kinds of lateral textures, depending on the coagulation conditions, have been observed in cellulose II and PPTA fibers. In addition, the LC phase fibers can display a pleated sheet morphology with a period of about 250 nm. Typical for the wet-spinning process are the finely distributed and highly elongated voids in the fiber, which give it a fibrillar morphology.

6. Mechanical Properties

The mechanical properties of LC phase solution spun fibers can be fully understood on the basis of a single-phase crystalline structure. Numerous experiments show that the continuous chain model provides a good interpretation of the mechanical properties. In this model the fiber properties are determined by the chain modulus (e_c), the shear modulus between adjacent chains (g), the corresponding critical shear stress (τ_c), the orientation distribution of the chain axes ($\langle \sin^2\Theta \rangle$), and the molecular weight distribution (MW). The fiber is modeled as a series arrangement of crystalline domains consisting of perfectly oriented chains. The fiber strain is brought about by elastic chain extension and by elastic shear of the chain axes towards the fiber axis. Both contributions form the equation for the initial modulus of the fiber (E) given by:

$$\frac{1}{E}=\frac{1}{e_c}+\frac{\langle \sin^2\theta \rangle}{2g}$$

where θ is the angle between the chain axis and the fiber axis. For a crack-free fiber the internal shear modulus is about equal to the torsion modulus of the filament.

High draw ratios in the spinning process improve the chain orientation resulting in a higher modulus. Other process conditions, such as hot drawing, can influence the molecular interactions to some degree, e.g., by improving the crystalline order, and hence may increase the shear modulus. Table 1 presents the significant data and the types of interchain bonds. Depending on the degree of chain orientation the tensile curves of polymer fibers below their glass transition temperature are characterized by a more or less pronounced yield at about 0.005–0.025 strain. This yield can be modeled using a critical resolved shear stress, τ_c, at which local loosening of the intermolecular bonds occurs, resulting in a permanent rotation and displacement of adjacent chains and, thus, in a permanent shear deformation of the domains. Hence, above the yield strain the tensile extension of the fiber is accompanied by plastic deformation. It has been shown that the compressive strength of polymer fibers is determined by the same critical shear stress leading to yielding in compression occurring at the same strain as in tensile deformation. Generally, the compressive strength of a crack free

Table 1

Summary of the basic elastic constants g and e_c for various organic polymer fibers, together with the highest filament values reached for the modulus (E) and the tensile strength (σ_b). Also shown are the average values for the creep compliance (CC) at 20 °C, which is the ratio of creep rate and load stress, and the character of the lateral interchain bonds.

Fiber	e_c (GPa)	g (GPa)	E (GPa)	σ_b (GPa)	CC (%/dec/GPa)	Interchain bond
UHMW PE	250–300	0.7–0.9	150	6.0	0.3	Van der Waals
PET	125	1.0–1.3	38	2.3	0.13	Van der Waals/dipolar
Cellulose II	88	2.0–5.0[a]	55	1.7	0.4[b]	Bi-directional hydrogen
Cellulose I[c]	136	1.5	110	–	–	Unidirectional hydrogen/Van der Waals
PPTA	240	1.5–2.7	140	4.5	0.02–0.04	Unidirectional hydrogen/Van der Waals
PBO	500	1.0–1.5	260	6.1	0.005	Van der Waals/dipolar
PIPD[d]	550	6.0	330	5.0	0.0035	Bi-directional hydrogen

[a] Depending on water content. [b] Conditioned at 65% RH. [c] native fiber. [d] PIPD or M5 is a fiber in a early stage of development so that further improvement of these data is expected.

Table 2

Some applications of LC spun high performance fibers.

Application	Used property	Which polymer
Brake linings	Modulus, T resistance	PPTA
Ballistics	Tenacity, modulus	PPTA, PBO, PIPD
Protective clothing	T resistance	PPTA, PIPD
Mechanical rubber goods	Tenacity, modulus	PPTA, Cellulose, PIPD
(e.g., conveyor belts)	Tenacity, modulus	PPTA, Cellulose, PIPD
Tires, pipes, and hoses	Tenacity, modulus	PPTA/Epoxy composites
Concrete reinforcement	Modulus	PPTA, PBO
Optical cables, fiber reinforced composites	Tenacity, modulus, T resistance	PPTA, PIPD

Table 3
Approximate production volumes of dry-jet wet-spun yarns from rigid polymers.

Brand	Polymer	Producer	Production tons/year
Twaron	PPTA	Teijin Twaron	12000
Kevlar	PPTA	DuPont	18000
Zylon	PBO	Toyobo	400
Cellulose[a]	Cellulose	Acordis	Laboratory scale
M5	PIPD	Magellan	Laboratory scale
Technora	p-aramid	Teijin	2000
Armos			Pilot plant scale
Rusar		Kamensk Chimvolokno	400

[a] New cellulose yarn from phosphoric acid solutions.

polymer fiber is about 0.3 times the shear modulus (this is supported by the yield model and experimental results). The highest experimental value for compressive strength of 1.7 GPa obtained so far is for the PIPD fiber, which is due to the network of hydrogen bonds that extends in two different directions normal to the chain. The hydrogen bond networks of cellulose II and PIPD are rather similar, which causes high shear moduli for both fibers. The intermolecular bonding also to a large extent determines the viscoelastic properties like the creep compliance of the polymer fibers as shown in Table 1. Strong LC phase-based fibers are obtained for large values of MW, τ_c, and a high degree of chain orientation, i.e., for high values of E and a more uniform structure of the fiber. The highly rigid nature of the polymer chain in PBO and PIPD, and the relatively strong interchain bonding, in particular in PIPD, ensure excellent thermal properties such as a low creep rate and long lifetimes at relatively high temperatures. It is noted that the creep rates for LC phase spun fibers (PPTA, PBO, and PIPD) are much lower than those for UHMW-PE and PET due to the stronger lateral interactions.

7. Applications

Due to their impressive mechanical and thermal properties, polymer fibers spun from LC solutions are extensively used for various reinforcement and structural applications. In general the advantage of such polymer fibers is that they have a high modulus and strength and are not brittle and have relatively good damage tolerance compared to, for example, carbon fiber and glass fiber. One of the largest applications of PPTA fibers is in the making of brake linings as a replacement of asbestos. Also PPTA aramid fibers are used for ballistics as the main component of bullet proof vests. Some additional applications are listed in Table 2. The various materials each have their own niche in the market and the estimated annual production is shown in Table 3.

See also: *Polymer Injection Molding: Multimaterial Molding*[#]; Polymer Fiber Processing: High-performance Fibers*; Processing Flexible Polymers to High Performance Fibers*; Main-chain Liquid Crystalline Polymers for Fibers[#]

Bibliography

Baltussen J J M, Northolt M G 1999 The stress and sonic modulus versus strain curve of polymer fibres with yield. *Polymer* **40**, 6113–24

Collyer A A 1992 *Liquid Crystal Polymers: From Structures to Applications*. Elsevier, London

Doi M, Edwards S F 1986 *The Theory of Polymer Dynamics*. Clarendon, Oxford

Kwolek S L 1972 *US Pat*. 3 671 542

Maier W, Saupe A 1959 Eine einfache molekular-statistiche Theorie der nematischen kristalflüssigen Phase. Teil I. *Z. Naturforsch*. **149**, 882–9

Northolt M G, Baltussen J J M 2001 The tensile and compressive deformation of polymer and carbon fibres. *J. Appl. Polym. Sci*. (in press)

Onsager L 1949 The effect of shape on the interaction of colloidal particles. *Ann. N.Y. Acad. Sci* **51**, 627

Robinson C, Ward J C, Beevers R B 1958 *Disc. Faraday Soc*. **25**, 29–42

S. J. Picken[a], H. Boerstoel[b] and M. G. Northolt[c]
[a] *Delft University of Technology, The Netherlands*
[b] *Teijin Twaron, Arnhem, The Netherlands*
[c] *AKZO Nobel, Arnhem, The Netherlands*

Pulp and Paper: Nonfibrous Components

Paper and paperboard consist of mainly of cellulosic fibers. In order to increase the quality and properties of these products, nonfibrous components are added. The types of materials and steps in the process are discussed in this article.

1. Wet End Addition

This addition or loading occurs when there is at least 60% fiber and before the formation of the paper or paperboard. Loading, a traditional term, means the incorporation of inorganic materials in the fibrous web to improve quality. The word "load" is used because a constant basis weight of the paper product is maintained during the addition of 1–40% inorganic material that has a higher density (2600–4500 kg m^{-3}) with a similar reduction of the less dense cellulose fibers (1500 kgm^{-3}). A web consisting of only fibrous materials will often have a discontinuous surface with minute hills, valleys, and randomly distributed voids. The objective is to close or plug the spaces between the fibers in the web of the paper product with nonfibrous materials. Another term used for the inorganic additives is fillers, essentially inorganic materials also called pigments. Different types of clay, calcium carbonate, titanium dioxide, talc, and specially made proprietary materials are examples of wet end fillers.

1.1 Clay

Kaolin is the most important clay material to improve surface smoothness for better print quality. Clay materials are natural, earthy, fine-grained, and plate-like materials, and the terms clay and kaolin are used interchangeably. Chemically, clays are hydrous aluminum silicates with an approximate composition of Al_2O_3–$2SiO_2$–$2H_2O$. At least 12 different classes of clays are available, which can be supplied in either a dry powder or a dispersed slurry water solution. In order to obtain maximum efficiency of clay, each particle must be completely separate or dispersed to ensure optimum flow properties. In its natural state, clay is acid and particles are attracted to each other (flocculated).

1.2 Calcium Carbonate

This additive increases brightness, ash, and opacity. Calcium carbonate can be divided into two general classifications: naturally occurring products such as limestone, granite, and chalk processed by wet or dry grinding; and those chemically produced by precipitation. Within these two groups are a wide variety of subgroups based on particle size, shape, and distribution. Lime is converted to calcium hydroxide which is reacted with carbon dioxide to form calcium carbonate and water. A number of on-site calcium carbonate plants have been constructed.

1.3 Titanium Dioxide

Titanium dioxide pigments are finely divided white powders which are chemically inert or unreactive, in contrast to all commonly used materials for paper filling or coating systems, and are used to increase opacity. Titanium dioxide is classified based on its crystalline arrangement as either anatase or rutile. The high light reflectivity, low light absorption and small particle size make this pigment ideal for obtaining high opacity. In their finely divided form, these pigments are very intensely white. With this characteristic, even at low concentration levels, titanium dioxide pigments are a major contributor to the optical performance of brightness and opacity of paper. If only brightness is needed, calcium carbonate is a more cost-effective pigment.

1.4 Talc

Talc is a silica combined with magnesium salts, being mined and processed similarly to clay, and used to decrease the amount of "pitch" in pulp.

1.5 Retention Aids

Another series of chemicals called retention aids are added at the wet end to ensure efficient capture of pigments during the paper-forming process. These materials are pigment specific and have a very complex chemistry.

2. Surface Treatment

Surface treatment of paper or paperboard is a general term used to describe the application of adhesives, functional products, pigments, and their combinations to the already formed fibrous web. Surface sizing and coating are designations that are frequently chosen to separate arbitrarily certain processes in surface treating.

3. Surface Sizing

Sizing occurs on the paper machine using specially designed equipment. The application of the treatment may consist of one material, usually a starch solution, or a combination of adhesives, pigments, and functional products. Again, the objective is to improve the appearance and print properties of the paper product. An example of a similar process would be the priming of bare wood.

4. Coating

This process can be accomplished either on or off the paper machine A coating formula or recipe contains three major components.

4.1 Coating Pigments

These are similar to the pigments previously described. Pigments are 70–80% of the coating formula. Properties affected by adding pigments are color, brightness, opacity, paper and ink gloss, smoothness, and holdout.

4.2 Adhesives

Also termed binders, adhesives are necessary to bond the coating to the paper/board and to the pigments. Normally, a coating formula contains 10–20% adhesives. There are two classifications of adhesives, natural and synthetic.

Protein and starch are examples of natural adhesives. Several types of industrial-grade soy protein are available. These are classified as unmodified (unhydrolyzed) and modified (hydrolyzed) types of high-, medium-, low-, and extra-low viscosity grades. The source of soy protein is defatted soybean flakes. Starch is a high molecular weight polymer, with many sources, including corn, wheat, potato, and waxy maize. Starches can be chemically modified for specific uses in coatings, surface sizing, or at the wet end of a paper machine. Starch must be gelatinized to be an adhesive, which is accomplished by cooking a starch–water mixture.

Synthetic coating adhesives are latexes or poly(vinyl alcohol). Latex is a synthetic resin emulsion, a dispersion of very small (0.1 μm) water-insoluble particles held in a water suspension by a balance of surface-active agents (surfactants or soaps). Poly(vinyl alcohol) is an unusually strong resin and provides excellent adhesive strength in binding pigments, cellulose, and other change particles. Poly(vinyl alcohol) is sold commercially as a powder or granular material that is white to cream in color. When dissolved in water, cast, and dried, a tough clear colorless film is formed. Poly(vinyl alcohol) is supplied in a variety of grades, which are dependent on the size and weight of the monomer. The type and level of adhesives affect not only the final strength but also brightness, opacity, paper and ink gloss, smoothness, ink and water absorption, flow properties, ease of drying, and coater runnability. The major pigment binding adhesives for the paper industry are polystyrene, a hard, inflexible thermoplastic polymer, and butadiene, a soft, flexible elastomer.

4.3 Additives

There are a number of materials in this classification which give specific properties to the coating formula. The amount added is usually 0.2–2% based on total pigment. Discussed below are some of the major additives.

(*a*) *Biocides.* Deterioration or spoilage of paper coating formulations and ingredients is a problem. Spoilage may show up as a loss in viscosity, change in pH, development of odor, or change in color. These conditions result in production problems, coating loss, and/or excessive losses to the sewer. Even though bacteria and fungi involved in spoilage are microscopic, their activities are readily visible in the final product.

(*b*) *Dispersants.* These chemicals are used to overcome the natural mutual attraction of pigment particles. If the particles are not separated, poor coater runnability develops. The separation is done through mechanical agitation or by chemical means. Slurry pigments have dispersants added by the suppliers. When using dry pigments in a formula, additional dispersant is required under high agitation conditions. Too much or too little dispersant will affect the coating viscosity.

(*c*) *Dyes.* Color changes are made through these additives, including fluorescent dyes. The addition of fluorescents improves the whiteness and blue-white hue of coated paper.

(*d*) *Flow modifiers.* Either natural or synthetic chemicals are used to modify the viscosity of coatings during the make-down, pumping, and application processes. A knowledge of the types of flow modifiers and how and where to add them to a coating mix is critical to the final quality of the coated sheet. Some examples of flow modifier problems include excessive dispersant in pigments or coating mix, not enough defoamer to minimize air entrapment (foam), and excessive caustic that causes pigment shock.

(*e*) *Lubricants.* There is no simple definition for the numerous properties and a function of a lubricant, the need of which is still debated. To impart lubrication to the wet coating, the following compounds may be used: soluble soaps (ammonium stearate and sodium sterate) to increase the viscosity of starch coatings; insoluble soaps (aluminum stearate and calcium stearate), which have a small effect on starch coating viscosity; polyethylene emulsion; or emulsion of various animal fats.

(*f*) *Insolubilizers.* Water resistance of paper coatings is very important in the printing process and the addition of an insolublizer imparts this property. The need for water resistance is determined by using a mixture of water and alcohol (fountain solution) in the printing process.

(*g*) *pH controls*. Materials can be acidic, neutral, or alkaline, of which pH is a measure. pH shock occurs when a strong acid or alkaline material is added to a coating mix in the wrong sequence. The result is a thick, lumpy coating similar to cottage cheese.

(*h*) *Plastic pigments*. A variety of synthetic polymers can be used as plastic pigments in place of minerals. The main criteria are that the plastic pigment particles be non-film-forming and remain discrete under conditions of application, drying, and finishing. They also should be insoluble in the binders used. Since plastic pigments are completely synthetic, their size, shape, mass, electrical charge, surface activity, compatibility, and fusion temperature can be well controlled.

(*i*) *Others*. Other materials can be added in small amounts. The types and amounts are proprietary and provide distinctive coated paper properties.

5. Summary

Paper and paperboard consist mainly of cellulosic fibers. In order to enhance the quality and properties of these products, nonfibrous components are added. Nonfibrous materials that are added to paper, paperboard, and coatings are dynamic as a result of ongoing research, specific applications, and customer demands. This article only identifies the major components relating to paper and paperboard coating surface and printing improvement.

See also: Paper Products: Container Board#; Paper Products: Food Packages#; Paper Products: Security Applications#

Bibliography

Hagemeyer R W (ed.) 1961 *Synthetic and Protein Adhesives for Paper Coating*, Monograph Series No 22. Technical Association Pulp and Paper Industry, New York

Hall H R (ed.) 1963 *Paper Coating Additives*, Monograph Series No 25. Technical Association Pulp and Paper Industry, New York

Libby C E (ed.) 1962 *Pulp and Paper Science and Technology Volume II: Paper*. McGraw-Hill, New York

Silvernail L H, Bain W M 1976 *Paper Coating Pigments*, Monograph Series No 38. Technical Association Pulp and Paper Industry, New York

D. P. Rawski
Port Edwards, Wisconsin, USA

Radiation Effects in Carbon–Carbon Composites

Carbon–carbon (C/C) composite materials are a class of high-temperature materials that consist of a carbon or graphite matrix, reinforced with carbon or graphite fibers (see *Carbon–Carbon Composites**, where this class of material is reviewed in detail). C/C materials that are used in nuclear environments are typically manufactured from graphitic fibers, and exhibit graphitic matrices by virtue of the matrix precursors and high heat treatment temperatures employed in their manufacture. Potential nuclear applications for C/C composites include high-temperature control rods for advanced, high-temperature, gas-cooled reactors (Burchell 1999), and the plasma-facing components of a fusion reactor (Snead 1999). Owing to its use as a moderator in nuclear fission reactors, much is known about the irradiation behavior of graphite (see *Nuclear Graphite and Radiation Effects*#). In contrast, comparatively little is known about irradiation damage effects in C/C composites (Burchell 1997).

1. *Carbon–Carbon Composites as Fusion Reactor Plasma-facing Materials*

In a fusion reactor, light elements are brought together at enormous temperatures and "fuse" to form a heavier element, energy being released in the process. There are many reactions that could be utilized in a fusion reactor. However, the fusion reaction that has been most widely studied is the deuterium–tritium reaction, which involves two isotopes of hydrogen:

$$_1\mathrm{H}^2 + _1\mathrm{H}^3 \rightarrow _2\mathrm{He}^4 + \text{neutron} + \text{energy}$$

In a fusion reactor, the hot plasma (ionized gas) is magnetically confined within a toroidal-shaped vacuum vessel. The plasma temperature required to sustain the fusion reaction is $> 50 \times 10^6$ K (> 10 keV). If the plasma were allowed to contact the vacuum vessel structural material, the high atomic number elements vaporized by the intense heat of the plasma would enter the plasma and create plasma instability. Consequently, it is necessary to protect or "armor" the inner walls of the vessel with a material of low atomic number, or a highly refractory material. The vessel armor is often referred to as a "plasma-facing component" (PFC) and the materials used are called "plasma-facing materials." The tokamak concept is one of the more popular of the highly specialized designs of fusion reactors (see *Fusion Reactors (Magnetically Confined)—Tokamaks: Materials*#). A tokamak utilizes a toroidal magnetic field confinement system in which the plasma ions are confined within, and coupled to, continuous magnetic field lines that travel helically through the toroidal vacuum vessel. Carbon–carbon composites have been widely used as tokamak plasma-facing materials, and are under consideration for use in "next-generation" tokamaks such as ITER (Barabash *et al.* 1998). The interactions that occur between the plasma-facing material and the plasma are particularly severe, and can result in failure due to high thermal stress, melting, sublimation, excessive surface erosion, and neutron damage. The effect of this latter phenomenon on C/C composites is the subject of this article.

2. *Neutron Irradiation Damage Mechanism*

In the graphite crystal, carbon atoms are covalently bound in a hexagonal array with a binding energy in the crystal lattice of about 7 eV. The lattice basal planes are stacked upon one another in an ABAB … stacking sequence with an optimal separation of 0.335 nm. The bonding between the basal planes is of the weak attractive type (Van der Waals). McEnaney (1999) gives a full account of the structure and bonding in carbon materials. Energetic particles, such as electron, ions, or fast neutrons, can displace carbon atoms from their equilibrium lattice positions. The energy required to displace a carbon atom is in the range 24–60 eV (Kelly 1981, Burchell 1996a). The displaced carbon atoms, referred to as primary knock-on carbon atoms (PKAs), may displace further carbon atoms themselves, through secondary collisions with the carbon atoms in the graphite lattice in a cascade effect. The cascade carbon atoms are referred to as secondary knock-on carbon atoms (SKAs). The displaced SKAs tend to be clustered in small groups of 5–10 atoms and are generally treated as if they occur randomly. The total number of displaced carbon atoms produced depends upon the energy of the PKA, which in turn is a function of the neutron energy spectrum and the neutron flux. Displaced carbon atoms recoil through the crystal lattice, displacing other carbon atoms and creating vacant lattice sites. A significant fraction of the PKAs and SKAs do not remain displaced, but diffuse between the graphite crystal layer planes and recombine with lattice vacancies. Others will coalesce to form C_2, C_3, or C_4 linear molecules that in turn form the nucleus of a dislocation loop—essentially a new graphite plane. Adjacent lattice vacancies in the same graphitic layer are believed to collapse parallel to the layers, thus forming sinks for other vacancies, and can no longer recombine and annihilate interstitial defects.

---

* Cross references marked by an asterisk are included in this volume.

# Cross references marked by a hash can be found by consulting the Encyclopedia of Materials: Science and Technology.

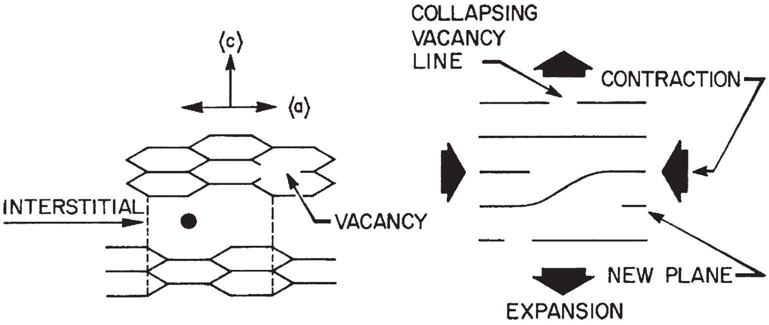


Figure 1
The radiation damage mechanism in graphite showing the induced crystal strains.

Figure 1 is a schematic representation of the neutron damage mechanism in graphite.

3. *Irradiation Induced Dimensional Changes*

Dimensional changes occur in the graphite crystallites because of the carbon atom displacements described above. Interstitial defects will cause the crystallites to expand perpendicularly to the basal planes (**c**-direction in Fig. 1), and relaxation in the plane due to the coalescence of vacancies will cause shrinkage parallel to the basal planes (**a**-direction in Fig. 1). The dimensional changes can be very large, as demonstrated in studies on well-ordered graphite materials, such as pyrolytic graphite (Price 1974). Typically, the dimensional change (**a**-axis shrinkage and **c**-axis growth) in pyrolytic graphite and graphite single crystals increases linearly with dose, but decreases with increasing graphitization temperature. The effect of graphitization temperature can be attributed to thermally induced improvements in crystal perfection, which reduces the number of vacancy trapping sites in the lattice.

Radiation damage in C/C composites has been recently reviewed and materials property data for fusion reactor plasma-facing C/C composites reported (Burchell 1996a, Burchell and Oku 1994). Two-directional (2D) C/C composites have been reported to exhibit excessive swelling perpendicular to the fabric layers (22–37% growth), and shrinkage in directions parallel to the fabric layers (typically 3–19%), on irradiation at 400 °C and to a dose of approximately 4 displacements per atom (dpa) (Burchell *et al.* 1991). C/C composites are an infinitely variable family of materials, and processing and design variables influence their properties and behavior. In order to select a C/C composite for a PFC application in which neutron damage will be significant, we must understand the effect of processing variables on neutron irradiation-induced structure and property changes. Burchell *et al.* (1992) irradiated 1D, 2D, and 3D C/C composites at 600 °C and to damage doses up to 1.5 dpa. Three-directional C/C composites were shown to exhibit more isotropic dimensional changes than either 1D or 2D C/C composite materials. Pitch fiber composites were shown to be more dimensionally stable than polyacrylonitrile (PAN) fiber composites, and high final heat-treatment (graphitization) temperatures were found to be beneficial. Similar trends were reported by Bonal and Wu (1996).

Burchell *et al.* (1996b) reported dimensional, strength, and thermal conductivity changes for so-called 1D, 2D, and 3D C/C composites irradiated to a peak neutron dose of 4.7 dpa at an irradiation temperature of 600 °C. Neutron irradiation-induced dimensional changes for three of the PAN fiber composite materials are summarized in Fig. 2. Specimen length and diameter changes are shown for a unidirectional fiber composite (UFC), a random fiber composite (RFC), and a 3D composite (223). The 1D composite exhibited extremely anisotropic dimensional changes, undergoing rapid shrinkage in the fiber-axis direction (length). In the direction perpendicular to the fiber axis (diameter), the composite shrank at a dose <1 dpa, followed by a reversal to expansion, returning to zero-dimensional change at approximately 2.5 dpa and continuing to expand at an increasing rate. A similar trend was observed for the 2D composites, as typified by the RFC material (Fig. 2), where the fiber axis is in the specimen diameter direction. In the off-axis direction (length), the composite exhibited a slight contraction followed by expansion, returning to the original length at approximately 2 dpa. The diameter (fiber-axis) directions of the RFC specimens exhibited shrinkage, although the magnitude of the shrinkage was much less for the same damage dose than in the case of the 1D composite. In contrast to the 1D and 2D composites, the 3D PAN fiber composite exhibited isotropic behavior at damage doses up to approximately 2 dpa. At doses exceeding 2 dpa, the composite **z**-direction (specimen length) continued to display shrinkage, whereas the fiber **x-y** (specimen

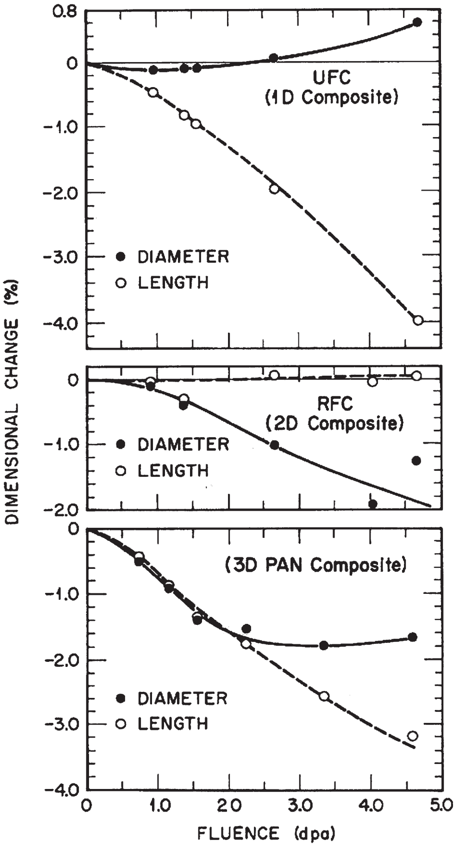


Figure 2
Neutron irradiation-induced dimensional changes of carbon/carbon composites irradiated at 600 °C.

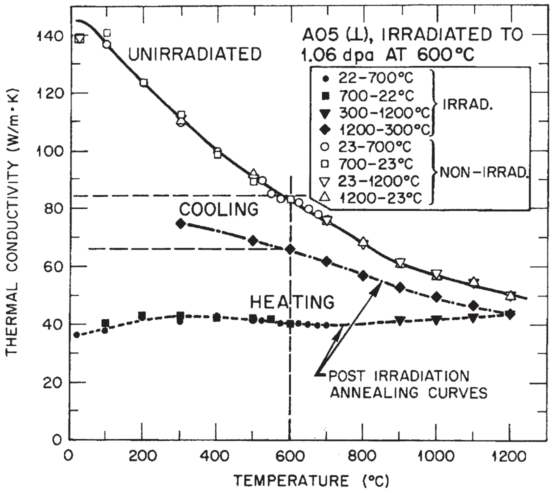


Figure 3
The temperature dependence of thermal conductivity for non-irradiated, irradiated, and annealed carbon/carbon composite material, AO-5.

diameter) direction exhibited reversal and slight growth. Irradiation-induced dimensional changes in C/C composites has recently been explained in terms of the following (Burchell 1996b):

(i) an understanding of the graphite single crystal dimensional changes;

(ii) a microstructure model of the carbon fibers used in the composite (PAN fibers in all three materials in Fig. 2); and

(iii) a knowledge of the composite architecture and the interactions of the fiber, fiber bundle, matrix, and porosity in the composite.

4. The Effects of Neutron Irradiation on Thermal Conductivity

Carbon and graphite are phonon conductors of heat. The thermal conductivity increases with increasing temperature from absolute zero, and peaks at a temperature typically close to room temperature, due to the opposing effects of the increasing specific heat and decreasing phonon mean free path – resulting from thermally induced phonon self-scattering (Umklapp scattering). The thermal conductivity decreases with further increases in temperature due to the dominating effect Umklapp scattering (Fig. 3). Phonon scattering also occurs from microstructural defects in the materials. Therefore, more graphitic materials exhibit greater thermal conductivity, and the peak in the thermal conductivity–temperature plot is sharper and shifts to lower temperatures. The mechanism of neutron irradiation-induced degradation of thermal conductivity in graphite has been extensively studied and reviewed (Burchell 1996a; Kelly 1981), and data for several C/C composites have been reported by Snead and Burchell (1995), Wu *et al.* (1994) and Eto *et al.* (1998).

The thermal conductivity is rapidly reduced with increasing neutron damage dose, due to the introduction of additional lattice defects, such as interstitial clusters, vacant lattice sites (singles, pairs or small groups of carbon atoms), and basal plane vacancy loops. Figure 3 illustrates the reduction of thermal conductivity with neutron damage dose. The lower curve in Fig. 3 illustrates the temperature dependence of the thermal conductivity of C/C composite A05 after neutron irradiation at 600 °C to ~1 dpa. At the irradiation temperature the thermal conductivity was reduced from $>80\,\mathrm{W m^{-1} K^{-1}}$ to $\sim 40\,\mathrm{W m^{-1} K^{-1}}$. Some fraction of the thermal conductivity may be recovered by thermally annealing the irradiation-induced defects, by heating the material above its irradiation temperature. This phenomenon may be seen in the "cooling" curve of Fig. 3. On cooling from the annealing temperature of 1200 °C to the irradiation temperature (600 °C), the thermal conductivity recovers to 67 $\mathrm{W m^{-1} K^{-1}}$. A greater recovery of thermal conductivity can be expected with increasing annealing temperature. The reduction in thermal conductivity due to irradiation damage is sensitive to dose and temperature. At any given dose and temperature, a saturation limit will be reached that is not exceeded until the composite undergoes gross structural changes at high damage doses. The saturation level of thermal conductivity will be lower, and will be attained more rapidly, at low irradiation temperatures.

5. *Conclusions*

Carbon–carbon composites are currently used in experimental fusion reactors, and are being considered for specialized high temperature gas-cooled fission reactors. In these nuclear applications, the material will be subjected to fast neutrons which result in structural damage and physical property degradation. Although relatively little work has been done to elucidate the effects of neutron damage in C/C composites, a vast body of literature exists for neutron damage in graphite. The data from the earlier studies on graphite has been used successfully to interpret new data on the effects of neutron damage in C/C composites.

See also: Radiation Damage Theory#; Nuclear Reactor Materials: Irradiation Effects#; Nuclear Reactors: Moderator and Reflector Materials#; Nuclear Reactors: Shielding Materials#

Bibliography

Barabash V, Akiba M, Bonal J P, Federici G, Matera R, Nakamura K, Pacher H D, Rödig M, Vieider G, Wu C H 1998 Carbon fiber composites application in ITER plasma facing components. *J. Nucl. Mater*. **258–63**, 149–59

Bonal J P, Wu C H 1996 Neutron irradiation effects on the thermal conductivity and the dimensional stability of carbon fiber composites. *Phys. Scr*. **T64**, 26–35

Burchell T D 1996a Radiation damage in carbon materials. In: *Physical Processes of the Interaction of Fusion Plasmas with Solids*. Academic Press, San Diego, CA, pp. 341–84

Burchell T D 1996b Radiation damage in carbon–carbon composites: structure and property effects. *Phys. Scr*. **T64**, 17–25

Burchell T D 1997 Radiation effects in graphite and carbon-based materials. *MRS Bull*. **XXII**, 29–35

Burchell T D 1999 Fission reactor applications of carbon. In: Burchell T D (ed.) *Carbon Materials for Advanced Technologies*. Elsevier, Oxford, pp. 429–84

Burchell T D, Eatherly W P, Hollenburg G H, Slagle O D, Watson R D 1991 The effects of neutron irradiation on the structure of carbon-carbon composites. In: *Proc. 20th Biennial Conf. Carbon*. American Carbon Society, pp. 598–9

Burchell T D, Eatherly W P, Robbins J M, Strizak J P 1992 The effect of neutron irradiation on the structure and properties of carbon–carbon composite materials. *J. Nucl. Mater*. **191–94**, 295–9

Burchell T D, Oku T 1994 Materials property data for fusion reactor plasma facing carbon–carbon composites. In: Janev R K (ed.) *Nuclear Fusion—Atomic and Plasma-Materials Interactions Data for Fusion*, **5** *Supplement to Nuclear Fusion*. IAEA, Vienna, pp. 77–128

Eto M, Baba S, Ishihara M, Ugachi H 1998 High energy neutron and charged particle irradiation effects on thermomechanical properties of carbon–carbon composites for divertor applications. *J. Nucl. Mater*. **258–63**, 843–7

Kelly B T 1981 *Physics of Graphite*. Applied Science, London

McEnaney B 1999 Structure and bonding in carbon materials. In: Burchell T D (ed.) *Carbon Materials for Advanced Technologies*. Elsevier Science, Oxford, pp. 1–33

Price R J 1974 High-temperature neutron irradiation of highly oriented carbons and graphites. *CARBON* **12**, 159–69

Snead L L 1999 Fusion energy applications. In: Burchell T D (ed.) *Carbon Materials for Advanced Technologies*. Elsevier, Oxford, pp. 389–427

Snead L L, Burchell T D 1995 Thermal conductivity degradation of graphites due to neutron irradiation at low temperature. *J. Nucl. Mater*. **224**, 222–9

Wu C H, Bonal J P, Thiele B 1994 Thermal conductivity changes in graphite and carbon/carbon fiber materials induced by low neutron damages. *J. Nucl. Mater*. **212–15**, 1168–73

T. D. Burchell
Oak Ridge National Laboratory, Tennessee, USA

Reaction Forming

The shrinkage upon firing of a porous ceramic-bearing preform can be reduced by increasing the solid volume within the preform prior to or during sintering. This may be accomplished by infiltrating a fluid (liquid or gas) into the porous preform and then converting the fluid into a solid by solidification, condensation, and/or chemical reaction with the preform. Approaches based on solidification or condensation will not be reviewed here. The purpose of the present contribution is to describe recent developments in processes that utilize oxidation reactions between an infiltrant (gas or liquid) and a shaped preform to produce near net-shaped, ceramic-bearing components. Two of these approaches yield all-ceramic bodies by the reactive infiltration of O_2(g) into shaped metal-bearing preforms: the reaction bonded aluminum oxide (RBAO) and volume identical metal oxidation (VIMOX) methods. In two other processes, ceramic/metal (intermetallic) composites are fabricated by the reactive infiltration of molten metals into porous, oxide-bearing preforms: the infiltration alumina aluminide alloy (i-3A) and displacive compensation of porosity (DCP) methods. In these processes the oxygen source is a solid oxide within the solid preform.

1. *Near Net-shaped Ceramics from Formable, Metal-bearing Preforms: The RBAO and VIMOX Methods*

Oxidation-induced changes in volume are utilized in the patented RBAO (Claussen 1992) and VIMOX (Sandhage 1997) methods to convert shaped metal-bearing precursors into ceramic bodies of similar shape and dimensions. The ratio of the volume of oxide produced to the volume of metal consumed is often referred to as the Pilling–Bedworth ratio (PBR). PBR values for several metals are shown in Table 1.

For most metals, the PBR is in excess of unity. The increase in solid volume associated with the oxidation of such metals within a porous body can be used to offset the contraction associated with postoxidation sintering. Hence, porous shaped preforms comprising intimate mixtures of metal and ceramic can be converted into dense all-ceramic bodies that retain the preform shape and dimensions. For example, the oxidation of a porous mixture of $Al + Al_2O_3 + ZrO_2$ followed by sintering can yield Al_2O_3/ZrO_2 composites with relatively low net shrinkage (Wu *et al.* 1993, Holz *et al.* 1994). This is the basic premise for near net-shape processing by the RBAO method.

Unlike most metals, the alkali and alkaline earth metals tend to possess PBR values less than unity (see the right half of Table 1). Such oxidation-induced reductions in solid volume can be used to offset volume expansions associated with the oxidation of other elements and/or the formation of ceramic compounds. That is, a preform composed of an appropriate mixture of alkali or alkaline earth elements with other metallic and/or ceramic constituents can be converted by oxidation and further reaction into a near net-shaped monolithic ceramic or ceramic composite. For example, a dense, shaped mixture of $Ba + Sr + Al + Al_2O_3 + SiO_2$ can be oxidized and annealed to produce a dense body of celsian, $(Ba,Sr)Al_2Si_2O_8$, with little or no change in shape and dimensions (Sandhage *et al.* 2000, Viers and Sandhage 1999). This is the basic premise for near net-shape fabrication by the VIMOX process.

Table 1
Pilling–Bedworth ratio (PBR) for several elements.

Metal/oxide	PBR	Metal/oxide	PBR
Al/Al_2O_3	1.28	Mg/MgO	0.804
Si/SiO_2	2.14	Ca/CaO	0.638
Cr/Cr_2O_3	2.01	Sr/SrO	0.611
Ni/NiO	1.66	Ba/BaO	0.670
Ti/TiO_2	1.77	Li/Li_2O	0.567
Zr/ZrO_2	1.51	Na/Na_2O	0.541
Nb/Nb_2O_5	2.69	K/K_2O	0.474

For both the RBAO and VIMOX methods, preforms composed of fine, intimate mixtures of metal and ceramic are required in order to obtain desired oxidation and reaction kinetics, and to produce near net-shaped bodies without macrocracking or distortion (Wu *et al.* 1993, Holz *et al.* 1994, Sandhage *et al.* 2000, Viers and Sandhage 1999, Schmutzler *et al.* 1994). High-energy milling can be an effective means of producing such mixtures. However, proper selection of the milling intensity (e.g., rotation speed during attrition milling), milling time, ball-to-charge ratio, and milling liquid are required to produce high yields of fine, intimate metal/oxide mixtures (Wu *et al.* 1993, Holz *et al.* 1994, Sandhage *et al.* 2000, Viers and Sandhage 1999, Schmutzler *et al.* 1994, 1996). Incomplete milling will yield coarse, nonuniform mixtures, whereas excessive milling may result in undesired mechanochemical reactions (e.g., the formation of brittle intermetallic compounds). Excessive mechanochemical reaction and oxidation of the metal-bearing powder mixture must be avoided during milling, so that a sufficient amount of ductile metal is retained to allow for: (i) uniform compaction and forming into a green body of desired shape; and (ii) small net volume changes upon conversion of the green body into the desired ceramic phase(s) (Holz *et al.* 1994, Sandhage *et al.* 2000, Viers and Sandhage 1999). Excessive hydration of the ceramic constituents during milling must also be avoided in order to reduce the risk of bloating during subsequent firing (Holz *et al.* 1994). Finally, owing to the inevitable wear that occurs during high-energy milling, the compositions of the milling media and milling vessel must be properly chosen to avoid degradation of the performance of the final fired components (Wu *et al.* 1993, Holz *et al.* 1994, Sandhage *et al.* 2000, Schmutzler *et al.* 1994).

Due to the presence of appreciable ductile metal in the milled powders (usually >30 vol.%), the green strengths of compacted RBAO and VIMOX mixtures can be appreciably higher than for conventional ceramic green bodies; that is, green strengths of 20–50 MPa are typically achieved in pressed metal-bearing RBAO and VIMOX preforms (Wu *et al.* 1993, Holz *et al.* 1994, Sandhage *et al.* 2000). In the RBAO process, compaction pressures of up to 500 MPa are commonly used to produce $Al–Al_2O_3$-bearing preforms with a sufficient amount of porosity (usually $\geqslant 25$ vol.%) so as to allow for a relatively high oxygen infiltration rate during oxidation and a good match between the oxidation-induced volume increase and the sintering-induced contraction. In the VIMOX process, preforms with densities of up to 93% have been produced using uniaxial compaction stresses as high as 1200 MPa. (Open porosity is not required in VIMOX preforms for near net-shape processing.) After compaction, the robust metal-bearing RBAO and VIMOX preforms can be readily machined into more complicated shapes (Wu *et al.*

1993, Viers and Sandhage 1999). Dense, metal-rich VIMOX preforms have also been formed into wires and tapes by drawing and rolling (Sandhage *et al.* 2000, Schmutzler *et al.* 1994, 1996).

Shaped metal-bearing VIMOX preforms can also be produced by melt infiltration. For example, dense Mg–Al_2O_3-bearing precursors to spinel ($MgAl_2O_4$) have been produced by the pressureless infiltration of molten magnesium into porous Al_2O_3 preforms followed by solidification (Sandhage *et al.* 2000, Kumar and Sandhage 1998). With careful tailoring of the porosity of the Al_2O_3 preform, the desired magnesium content for producing stoichiometric spinel can be achieved. The ductile, interconnected magnesium endows the infiltrated preform with a relatively high green strength (70–100 MPa). As for the milled, powder-derived preforms discussed above, melt-infiltrated preforms can be readily machined into more complex shapes (Kumar and Sandhage 1998).

Shaped, robust VIMOX and RBAO precursors can be converted into dense all-ceramic bodies with the use of appropriate oxidation and postoxidation heat treatments. The porous, nonprotective nature of the oxide scales formed on alkali or alkaline earth metals results in relatively rapid oxidation of VIMOX preforms at $\leqslant 500\,^\circ\mathrm{C}$ (Sandhage *et al.* 2000, Schmutzler *et al.* 1994, 1996). Although a dense, slow-growing layer of alumina tends to form on aluminum, a relatively high rate of oxidation of porous RBAO preforms can be achieved if the aluminum is in the form of fine (submicron) particles (Wu *et al.* 1993, Holz *et al.* 1994). (The protective nature of an alumina scale can also be reduced if the scale is allowed to react with other oxides, such as silica or baria, present in a preform (Sandhage *et al.* 2000, Viers and Sandhage 1999, Scheppokat *et al.* 1999)). During the oxidation of RBAO and VIMOX preforms, the following phenomena are of particular concern: (i) the formation of appreciable molten metal due to incomplete oxidation at subsolidus temperatures; and (ii) significant nonuniform thermal excursions (runaway oxidation) resulting from the exothermic nature of the oxidation reaction(s) (Holz *et al.* 1994, Sandhage *et al.* 2000, Schmutzler *et al.* 1994, Gaus *et al.* 1999). The generation of molten metal can lead to appreciable agglomeration and phase coarsening or to the loss of such metal from the preforms due to poor wetting (Holz *et al.* 1994, Sandhage *et al.* 2000, Schmutzler *et al.* 1994, Gaus *et al.* 1999). In order to maximize the extent of oxidation at temperatures where the metallic constituents remain solid, oxidation is conducted with modest rates of heating up to, and/or with isothermal anneals below, subsolidus temperatures. For example, 30–32.5 mm diameter × 35 mm tall Al–Al_2O_3-bearing cylinders have been completely oxidized by heating at 1–2 $^\circ\mathrm{C\,min}^{-1}$ to 1100 °C in air, in some cases with prolonged isothermal anneals at 400–475 °C (Wu *et al.* 1993, Gaus *et al.* 1999), and 10–16 mm lengths of sheathed Ba–Ce-bearing and Ba–Ti-bearing tapes have been completely oxidized (via oxygen migration from the exposed ends inward) within 4–10 h at 300 °C in oxygen (Sandhage *et al.* 2000, Schmutzler *et al.* 1994, 1996). Such modest rates of heating and low-temperature annealing also help to reduce the occurrence of thermal excursions within the preforms that lead to extensive macrocracking (Sandhage *et al.* 2000, Schmutzler *et al.* 1994, Gaus *et al.* 1999). Reductions in the oxygen partial pressure, and the use of gas mixtures containing helium (i.e., to enhance the rate of heat transfer through gas-filled pores), have also been effective means of reducing thermal excursions and associated macrocracking (Viers and Sandhage 1999, Gaus *et al.* 1999).

The oxide grains formed within RBAO and VIMOX preforms upon oxidation at modest subsolidus temperatures tend to be submicron in size (Wu *et al.* 1993, Sandhage *et al.* 2000). Such an intimate mixture of fine-grained oxides can undergo further reaction and sintering under modest annealing conditions (Sandhage *et al.* 2000, Schmutzler *et al.* 1996, Scheppokat *et al* .1999). For example, tapes of $BaTiO_3$ and $Ba(Ce,Nd)O_3$ with densities of $\geqslant 98\%$ have been produced from well-milled and oxidized preforms at peak temperatures of $\leqslant 1085\,^\circ\mathrm{C}$. Intentionally porous bodies of $Ca_{10}(PO_4)_6(OH)_2$ and $(Ba,Pb)TiO_3$ have been produced from well-milled and oxidized VIMOX preforms within times of only 1 h and 10 min, respectively, at peak temperatures of 1120–1150 °C. Dense ($\geqslant 95\%$) Al_2O_3/ZrO_2 composites have also been synthesized by firing fully-oxidized RBAO preforms for only 20 min at 1550 °C. Some care has to be taken, however, with such fast firing to avoid bloating. Bloating can occur if sintering at a specimen surface precedes complete removal of chemically-bound water (e.g., in the form of hydroxides) located deeper within the specimen. Preforms prepared with high compaction stresses from powders milled for extended times in water-bearing lubricants are particularly susceptible to such bloating (Holz *et al.* 1994).

With proper tailoring of the preform composition, phase content, and microstructure and with appropriate heat treatment conditions, easily-formed RBAO and VIMOX green bodies can be oxidized and converted into near net-shaped ceramics and ceramic composites with attractive properties. For example, dense (pressureless sintered) RBAO-derived Al_2O_3–ZrO_2 composites have been produced with fracture strengths and toughnesses of 700–800 MPa and 5–6 $\mathrm{MPa\,m}^{-2}$, respectively. By incorporating metal or ceramic additives with larger oxidation-induced volume changes than for aluminum (e.g., Si, SiC, Zr, Cr), nearly zero shrinkage has been achieved by the reaction bonding process (Wu *et al.* 1993, Scheppokat *et al.* 1999). For example, porous preforms of Al– Al_2O_3–SiC–ZrO_2 have been converted into 93–99% dense composites of $Al_6Si_2O_{13}$ (mullite) and ZrO_2 with net linear shrinkages of $\leqslant 1\%$ by slow heating in air (in some cases with several intermediate isothermal anneals) up to 1150–1200 °C, holding at this temperature for 15 h, and then sintering at 1550 °C (Scheppokat *et al.* 1999). Incomplete oxidation of the SiC has yielded $Al_6Si_2O_{13}/SiC/ZrO_2$-bearing composites with modest net shrinkages (e.g., $\leqslant 7\%$). The latter composites have exhibited fracture strength and toughness values of up to 610 MPa and 4.9 $MPa\,m^{-2}$, respectively.

A variety of functional, near net-shaped ceramics and ceramic composites have been fabricated by the VIMOX process for chemical, electronic, magnetic, gas-sensing, and biomedical applications. Such ceramics include refractory aluminates (e.g., $MgAl_2O_4$, $BaAl_2O_4$, $Ba(Sr)Al_2Si_2O_8$), dielectric or semiconducting titanates (e.g., $BaTiO_3$, $(Ba,Pb)TiO_3$), superconducting cuprates (e.g., $Bi_2Sr_2Ca_1Cu_2O_8$), magnetic ferrites (e.g., $BaFe_{12}O_{19}$), proton-conducting cerates (e.g., Nd_2O_3-doped $BaCeO_3$), water-sensitive chromates (e.g., TiO_2-doped $MgCr_2O_4$), and biocompatible phosphates (e.g., $Ca_{10}(PO_4)_6(OH)_2$). For example, machinable Ba(Sr)–Al–Al_2O_3–SiO_2-bearing preforms have been fully oxidized and converted into $Ba(Sr)Al_2Si_2O_8$ (celsian, a radome ceramic) bodies that retained the preform shapes with dimensional changes well under 1% (Sandhage *et al.* 2000, Viers and Sandhage 1999). Dense, machined Mg–Al_2O_3-bearing preforms have been fully oxidized and converted into dense bodies of spinel ($MgAl_2O_4$, a refractory ceramic) with dimensional changes of $\leqslant 0.6\%$ (Sandhage *et al.* 2000, Kumar and Sandhage 1998). Optical images of green-machined RBAO and VIMOX preforms, and corresponding fully-converted all-ceramic bodies, are shown in Fig. 1.

2. *Net-shaped Ceramic/Metal(Intermetallic) Composites from Porous Oxide-bearing Preforms: The i-3A and DCP Methods*

In the RBAO and VIMOX processes, a fluid oxidant (gaseous oxygen) is allowed to infiltrate into and react with shaped, metal-bearing preforms. A converse approach is to allow a fluid metal to infiltrate into and react with a shaped oxidant-bearing preform. The latter approach is the basis of two more recently developed, reactive infiltration techniques for fabricating dense, near net-shaped ceramic/intermetallic or ceramic/metal composites: the infiltration alumina aluminide alloy (i-3A) process and the displacive compensation of porosity (DCP) process.

A variety of high-melting, alumina–aluminide alloy composites can be fabricated at modest temperatures using the i-3A pressure infiltration process (Wagner *et al.* 1999, Schen *et al.* 1998). In this approach, a shaped, porous preform is first prepared by milling a mixture of a solid oxidant (e.g., TiO_2, Fe_2O_3, Nb_2O_5, ZrO_2) with a quantity of Al_2O_3. The powder mixture is compacted and formed into a shaped green body that is then partially sintered under modest firing conditions (e.g., 30 min at 1250°C). Such sintering produces a rigid (inert) Al_2O_3 network that endows the green body with reasonable mechanical stability during infiltration and reaction. The mechanical stability may be further enhanced by adding water glass or ductile metals to the oxide-rich preform. The partial sintering treatment is also used to adjust the pore fraction of the preform to a predetermined value. By carefully controlling the preform porosity along with the volume fractions of solid oxidant and alumina, it is possible to produce dense, near net-shaped composites composed of tailored amounts of alumina and an intermetallic compound (Wagner *et al.* 1999, Schen *et al.* 1998). Consider the fabrication of dense $Al_2O_3/TiAl_3$ and $Al_2O_3/TiAl$ composites via the following displacement reactions:

$$3TiO_2(s) + 13Al(1) \rightarrow 2Al_2O_3(s) + 3TiAl_3(s) \quad (1)$$

$$3TiO_2(s) + 7Al(1) \rightarrow 2Al_2O_3(s) + 3TiAl(s) \quad (2)$$

The product of Eqn. (1) consists of 30.8 vol.% Al_2O_3 with 69.2 vol.% $TiAl_3$. In order to convert a preform undergoing this reaction into a $TiAl_3$-bearing composite with a higher alumina content, the preform should contain excess (inert) alumina. For example, a 60 vol.% Al_2O_3–40 vol.% $TiAl_3$ composite can be produced by the infiltration and reaction of molten aluminum with a 54.2% dense preform comprising 68.3 vol.% Al_2O_3 and 31.7 vol.% TiO_2 (rutile). (This calculation assumes that infiltration is conducted at 850°C, where the density of Al(l) is 2.332 $g\,cm^{-3}$, and that the reaction proceeds along with the infiltration.) If the preform density and alumina phase content are chosen to be 46.8% and 61.3 vol.%, respectively, then Al(l) infiltration and reaction should yield a dense composite of 60 vol.% Al_2O_3–40 vol.% TiAl (via Eqn. (2)). Similar calculations can be conducted for the fabrication of other well-defined 3A composites.

To enhance the rate of Al(l) infiltration, elevated gas pressure and direct squeeze casting have been used in the i-3A process (Wagner *et al.* 1999). With gas pressure infiltration, a porous specimen is immersed under a molten aluminum bath and the pressure of an inert gas (e.g., argon) over the bath is then gradually increased to a modest value for a short time (e.g., 12 MPa for 10–20 min). After infiltration, the preform is raised out of the melt and cooled to room temperature. With squeeze casting, the preform is preheated (e.g., to 600 °C) and then quickly placed into a heated die (e.g., at 300–600 °C). Molten aluminum is then poured into the die and a ram is used to squeeze Al(l) into the porous preform. After solidification of the aluminum, the infiltrated preform and excess solid aluminum are removed from the die.

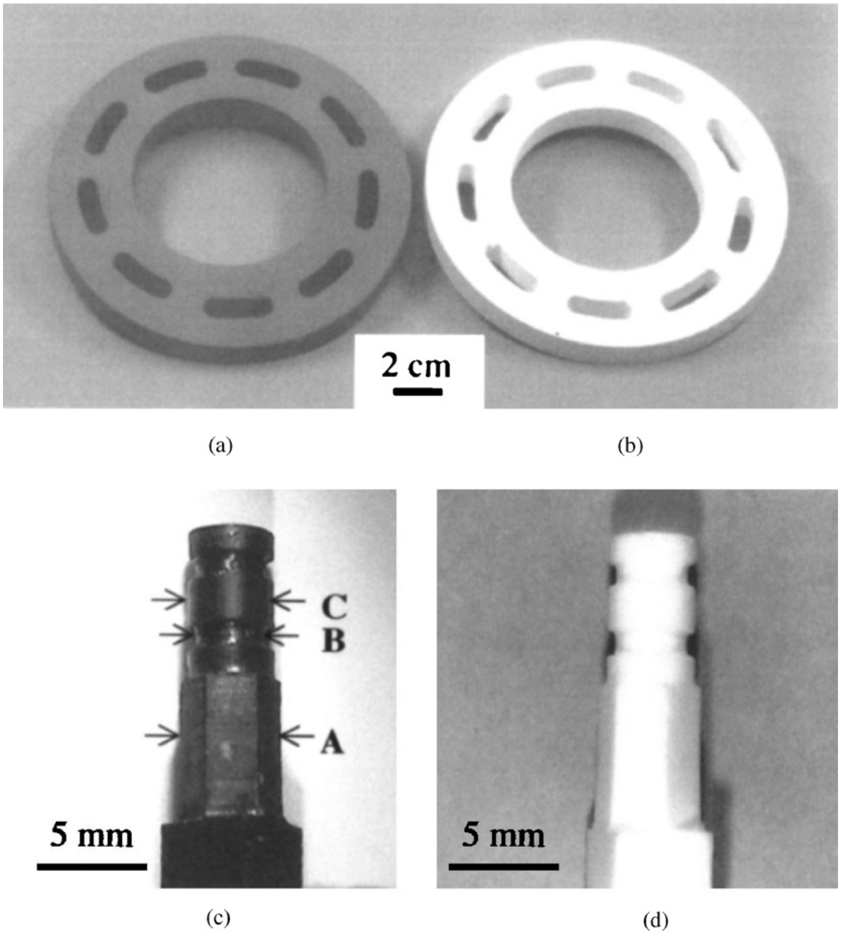


Figure 1
Optical images of: (a) a machined Al–Al_2O_3–ZrO_2-bearing green body; (b) a similar body after full conversion into a low-shrinkage Al_2O_3/ZrO_2 composite (RBAO process); (c) a Mg–Al_2O_3-bearing green body; and (d) the resulting $MgAl_2O_4$ body (VIMOX process (Kumar and Sandhage 1998)). Dimensional changes at the positions A, B, and C in Figs. (c) and (d) were $\leqslant 0.6\%$.

Although squeeze casting results in particularly rapid infiltration (i.e., 6 mm thick preforms can be completely infiltrated within a fraction of a second (Beyer *et al.* 2000)), additional processing must be undertaken to remove the excess solidified aluminum surrounding the specimen and possibly to complete the displacement reaction.

Displacement reactions between molten aluminum and solid oxidants tend to be relatively exothermic (e.g., the standard enthalpy per mole of reaction in Eqn. (1) is 1120 kJ at 850 °C). Such evolved heat can dramatically enhance the rate of reaction at modest applied temperatures. Indeed, dense, fully-reacted Al_2O_3–$TiAl_3$-bearing and Al_2O_3–$NbAl_3$–Nb_2Al-bearing composites have been produced by gas pressure infiltration of Al(l) into Al_2O_3/TiO_2 and $Al_2O_3/Nb_2O_5/NbAlO_4$ preforms, respectively, at peak pressures of 12 MPa for $\leqslant 20$ min at $\leqslant 850\,°C$ (Wagner *et al.* 1999, Schen *et al.* 1998). In the case of squeeze casting, where the residence time of molten aluminum in the preform can be very short ($\ll 1$ s), a secondary heat treatment at a modest temperature may be required to complete the reaction (e.g., 640 °C for 4 h to produces an Al_2O_3–$TiAl_3$-bearing composite (Wagner *et al.* 1999)). Because the net volume change upon completion of the i-3A displacement reaction is negative (e.g., -19.6% for the reaction in Eqn. (1) above), some internal porosity can develop if such a secondary heat treatment is conducted in the absence of excess aluminum surrounding the preform. Alumina-bearing composites reinforced with a variety of interpenetrating intermetallic compounds (e.g., $TiAl_3$, FeAl, $NbAl_3$, Nb_2Al) have been produced by the i-3A process.

The DCP process also involves the conversion of porous ceramic-bearing preforms into dense composites by the reactive infiltration of molten metals. However, unlike the i-3A method, the ceramic products of DCP-type reactions possess larger volumes than the ceramic reactants (Kumar and Sandhage 1999, Rogers *et al.* 1999). Such reaction-induced increases in ceramic volume can be used to fill the prior pore spaces (i.e., reaction-induced densification of the ceramic preform). The infiltration of shaped, rigid, partially-sintered preforms, followed by complete reaction at internal solid/liquid interfaces, can yield dense ceramic-rich composites that retain the shape and external dimensions of the starting preforms. Examples of two DCP-like reactions are:

$$3Mg(l) + Al_2O_3(s) \rightarrow 3MgO(s) + 2Al(l) \qquad (3)$$

$$3Mg(l) + 4Al_2O_3(s) \rightarrow 3MgAl_2O_4(s) + 2Al \qquad (4)$$

The reactions in Eqns.(3) and (4) are thermodynamically favored for Al–Mg melts containing more than about 0.3 at.% Mg and about 7.1 at.% Mg, respectively, at 1000 °C. Hence, as for the i-3A approach, these DCP reactions can be conducted with Al-rich melts. The volume of 3 moles of MgO is 31.9% larger than the volume of 1 mole of Al_2O_3, whereas the volume of 3 moles of $MgAl_2O_4$ is 16.6% larger than the volume of 4 moles of Al_2O_3. Consequently, a porous Al_2O_3-bearing preform can be converted into a MgO-bearing or $MgAl_2O_4$-bearing composite with a higher ceramic content than in the starting preform. Indeed, by adjusting the porosity and phase content of the DCP preforms along with the melt composition, fully-reacted ceramic/metal composites with a wide range of phase contents (i.e., from discontinuous to continuous metal reinforcements) and compositions can be directly produced. For example, 71%-dense Al_2O_3 preforms have been pressureless-infiltrated and reacted with pure Mg(l) at 1000 °C to produce composites of 86 vol.% MgO + 14 vol.% Mg–Al alloy; in this case, the metal phase was discontinuous, as demonstrated by electrical conductivity measurements (Kumar and Sandhage 1999, Rogers *et al.* 1999). Cocontinuous $MgAl_2O_4$/Fe–Ni–Al alloy-bearing composites (with 42–59 vol.% oxide) were fabricated by the pressureless infiltration and reaction of Mg–Al liquid at 900 °C with porous preforms comprising mixtures of Fe and $NiAl_2O_4$ (Rogers *et al.* 1999). Such DCP-derived composites were dense and retained the shapes and dimensions of the starting porous preforms to within a few percent (Kumar and Sandhage 1999, Rogers *et al.* 1999).

Near net-shaped oxide/metal and oxide/intermetallic composites with attractive, tailorable properties can be produced by the i-3A and DCP methods using low-temperature, low-cost casting technology. Lightweight $Al_2O_3/TiAl_3$ composites produced by the i-3A method have exhibited average strength and toughness values of 540 MPa and 8.6 $MPa\,m^{-2}$, respectively (Wagner *et al.* 1999, Schen *et al.* 1998). $MgAl_2O_4$/Fe–Ni–Al alloy-bearing composites produced by the DCP method have exhibited average strength and toughness values of 380 MPa and 11.3 $MPa\,m^{-2}$, respectively (Rogers *et al.* 1999). Composites that are electrically conductive (i.e., with a continuous metal or intermetallic phase) or electrically insulating (i.e., with a discontinuous metal phase) have also been fabricated (Wagner *et al.* 1999, Kumar and Sandhage 1999, Rogers *et al.* 1999).

Bibliography

Beyer P, Janssen B, Claussen N 2000 Synthesis of aluminide-alumina composites by reactive squeeze casting. *Adv. Eng. Mat.* **2**, 734–7

Claussen N 1991 Fabrication and properties of low-shrinkage reaction-bonded mullite. *J. Am. Ceram. Soc.* **74**, 2460–3

Claussen N 1992 *US Pat.* 5158916. Process of producing powder metallurgically produced ceramic formed body

Gaus S P, Sheedy P M, Caram H S, Chan H M, Harmer M P 1999 Controlled firing of reaction-bonded aluminum oxide (RBAO) ceramics: part II, experimental results. *J. Am. Ceram. Soc.* **82**, 909–15

Holz D, Wu S X, Scheppokat S, Claussen N 1994 Effect of processing parameters on phase and microstructure evolution in RBAO ceramics. *J. Am. Ceram. Soc.* **77**, 2509–17

Kumar P, Sandhage K H 1998 The fabrication of near net-shaped spinel bodies by the oxidative transformation of Mg/Al_2O_3 precursors. *J. Mater. Res.* **13**, 3423–35

Kumar P, Sandhage K H 1999 The displacive compensation of porosity (DCP) method for fabricating dense, shaped, high-ceramic-bearing bodies at modest temperatures. *J. Mater. Sci.* **34**, 5757–69

Rogers K A, Kumar P, Citak R, Sandhage K H 1999 Dense, shaped ceramic/metal composites at $\leqslant$1000 °C by the displacive compensation of porosity method. *J. Am. Ceram. Soc.* **82**, 757–60

Sandhage K H 1995 *US Pat.* 5447291

Sandhage K H, Allameh S M, Kumar P, Schmutzler H J, Viers D, Zhang X-D 2000 Near net-shaped, alkaline-earth-bearing ceramics for electronic and refractory applications via the oxidation of solid, metal-bearing precursors (the VIMOX process). *Mater. Manuf. Proc.* **15**, 1–28

Schen C, Dehm G, Kaplan W D, Wagner F, Claussen N 1998 Microstructure and phase evolution of niobium-aluminide-alumina composites prepared by melt-infiltration. *Phys. Stat. Sol.* (*a*) **166**, 241–55

Scheppokat S, Janssen R, Claussen N 1999 Phase development and shrinkage of reaction-bonded mullite composites with silicon carbide of different particle sizes. *J. Am. Ceram. Soc.* **82**, 319–24

Schmutzler H J, Antony M M, Sandhage K H 1994 A novel reaction path to $BaTiO_3$ by the oxidation of a solid metallic precursor. *J. Am. Ceram. Soc.* **77**, 721–9

Schmutzler H J, Sandhage K H, Nava J C 1996 The fabrication of dense, shaped barium cerate by the oxidation of solid metal-bearing precursors. *J. Am. Ceram. Soc.* **79**, 1575–84

Viers D S, Sandhage K H 1999 Near net-shaped (Ba,Sr)Al_2-Si_2O_8 bodies by the oxidation of machinable metal-bearing precursors. *J. Am. Ceram. Soc.* **82**, 249–52

Wagner F, Garcia D E, Krupp A, Claussen N 1999 Interpenetrating Al_2O_3-$TiAl_3$ alloys produced by reactive infiltration. *J. Eur. Ceram. Soc.* **19**, 2449–53

Wu S X, Holz D, Claussen N 1993 Mechanisms and kinetics of reaction-bonded aluminum oxide ceramics. *J. Am. Ceram. Soc.* **76**, 970–80

K. H. Sandhage[a] and N. Claussen[b]
[a] *Ohio State University, Columbus, Ohio, USA*
[b] *Technische Universitat Hamburg–Harburg, Hamburg, Germany*

Reactive Ceramic Nanocomposites with Organic and Bio-organic Phases

Ceramics and glasses can be modified into useful functional materials by doping with organic molecules, including proteins, via the sol–gel process. Taking into account the vastness of organic chemistry with its 23 million molecules and endless reactions, it becomes obvious that this new trend in materials chemistry ensures plenty of applications. This article provides some representative studies, mainly since the early 1980s, which marked the emergence of this field from obscurity into a self-contained domain of research and development (Avnir *et al.* 1998).

1. *General Aspects of Doping of Sol–Gel Materials with Organic and Bioorganic Molecules*

The inherent difficulty in combining organic chemistry with glass and ceramic chemistry which the ambient-temperature sol–gel process has successfully solved, is the temperature barrier. Organic compounds rarely survive temperatures higher than 200 °C and biomolecules far less. It is for this reason that over the millennia of the history of ceramics and glass production, one finds that the only additives, e.g., colorants, which could withstand the high temperatures were inorganic salts and other oxides. The sol–gel process has thus closed the ceramics–organic gap by utilizing it up to the porous xerogel stage.

Most of the inorganic oxides, MO_n (where M is a metal or semimetal, and n is not necessarily an integer) as well as many mixed oxides, have been used for this purpose (Brinker and Scherer 1990). Many studies concentrated on SiO_2 as a convenient model, although for most applications described below this need not be the optimal matrix, and one may also consider other oxides such as TiO_2, ZrO_2, Al_2O_3, composites such as SiO_2–ZrO_2, and organically modified matrices obtainable by either inorganic–organic copolymerizations or by the use of organic substituents on the inorganic precursor. The formation of the sol–gel materials is through the hydrolysis–polycondensation of monomers such as tetramethoxyorthosilicate (TMOS), but many other alkoxides, aryl oxides, and esters, chlorides and metal–nitrogen compounds can be used. The resulting materials—$(SiO_nH_m)_p$ in the case of, e.g., TMOS—are highly porous with pore size distributions centered typically around several tens of angstroms with surface areas of several hundreds of square meters per gram in most cases. The doping procedure is simple: the guest molecule is added to the polymerizing mixture either to the solution of monomers, or to a sol obtained by partial prehydrolysis and condensation. When the polycondensation is completed, the dopant molecules are entangled in the inorganic polymeric network. Intensive studies (Avnir 1995) of organic molecules thus entrapped in sol–gel matrices have provided the following generalizations.

- Practically any desired molecule can be entrapped in sol–gel matrices.
- The entrapped molecule usually retains its characteristic physical properties. Of the many consequent applications, optical applications are the most developed.
- The entrapped molecule usually retains its chemical properties. Chemical reactions and interactions are thus possible.

Since this article focuses on applications of *reactive* organically doped sol–gel silicas, it is in order to mention some of the other properties which have contributed to the attractiveness and versatility of this approach. These include chemical, photochemical, and electrochemical inertness, thermal stability of the matrix, the ability to induce electrical conductivity (Tsionsky *et al.* 1994), the controllability of matrix structural properties and the many ways available to modify chemically the matrix and its surface, the enhanced stabilization of the entrapped molecule (Negishi *et al.* 1994), the ability to employ the chromatographic properties of the matrix for enhanced selectivity and sensitivity of reactions with the dopant, the simplicity of the entrapment procedure which does not require chemical modification of the dopant, and the ability to obtain the doped sol–gel materials in any desired form, including powders, monoliths, films, fibers, stable sols and more.

Within the family of reactive doped sol–gel materials, this article concentrates on three major areas: organometallic catalysis, enzymatic reactivity, and sensing processes. Several other important classes, e.g., photochromic isomerizations (Levy *et al.* 1989) are not included for lack of space. The huge field of organically doped sol–gel optics (Klein 1994, Jorgensen and Reisfeld 1996) is also omitted, which includes useful materials such as light filters, luminescent materials, light guides, light collectors, dye-laser components, photochromic materials, hole-burning materials, optical memory materials, nonlinear optical materials, and liquid crystal displays (Levy *et al.* 1997). Information on such systems is given in *Ceramic Nanocomposites with Organic Phases, Optics of**.

2. Catalysis with Entrapped Organic Complexes

Sol–gel materials have a long history in serving catalysis. Celebrated examples are the pioneering studies of Teichner *et al.* (Bianchini *et al.* 1981) who used as catalysts the pure oxides (e.g., alumina), the use of mixed oxides, such as NiO–alumina–silica (Monaco and Ko 1998, Livage 1998), the use of sol–gel supports such as silica, alumina, and titania for the preparation and entrapment of dispersed metals (Gonzales *et al.* 1997), and the copolymerization of pure metal–alkoxides with metal alkoxide derivatives which carry as a substituent an organic metal ligand (Lorenz *et al.* 1997). This section concentrates on the direct physical entrapment approach, in contrast to catalysts which carry covalently bound organic moieties. Physical entrapment has enabled the conversion of a variety of soluble metal compounds into leach-proof recyclable catalysts that usually proved to be more stable under ambient conditions than their homogeneous analogues. These heterogeneous catalysts are highly selective, but are limited to reactions with substrates which can penetrate the small pore network of the sol–gel matrix to reach the catalyst.

A wide range of organometallic catalysts of differing natures can be entrapped in sol–gel materials. Representative examples include $RhCl_3$–$[(C_4H_9)_4N]Cl$, $RhCl_3$–$[(C_7H_{15})_4N]Br$, $RhCl_3$–$[(C_8H_{17})_3NMe]Cl$, $RhCl_3$–$[(C_{16}H_{33})P(C_4H_9)_3]Br$, $PtCl_4$–$[(C_8H_{17})_3NMe]Cl$ (Blum *et al.* 1996, Rosenfeld *et al.* 1996), $RuCl_2(PPh_3)_3$, $RhCl(PPh_3)_3$, $IrCl(CO)(PPh_3)_2$, η^5-$C_5H_5Rh(CO)_2$, η^5-$C_5H_5Rh(cod)$, $[Rh(CO)_2(\mu\text{-}Cl)]_2$, $[Rh(CO)PPh_3(\mu\text{-}Cl)]_2$, $[Rh(CO)(PPh_3)(\mu\text{-pyrazolate})]_2$ (Sertchook *et al.* 1996, Schumann *et al.* 1998), heterodinuclear complexes such as $Co_2Rh_2(CO)_{12}$ (Labroue and Poilblanc 1977), hydrophilic catalysts such as $RuCl_2[Ph_2P(3-C_6H_4SO_3Na)]_2 \cdot 4H_2O$, $RuCl[Ph_2P(3-C_6H_4SO_3Na)]_2 \cdot 4H_2O$ and $IrCl(CO)[Ph_2P(3-C_6H_4SO_3Na)]_2$, and more.

Representative examples for the applications of these catalysts are the hydrogen transfer processes. The first case is the disproportionation reaction of *vic*-dihydroarenes. In a typical application of the sol–gel methodology, the entrapped ion pair $[(C_8H_{17})_3NMe]^+[RhCl_4 \cdot nH_2O]^-$ affected a series of 1,4-dihydroarenes, yielding exclusively the expected disproportionation products in equimolar quantities, whereas under similar conditions, the nonimmobilized ion pair proved practically inactive (Rosenfeld *et al.* 1996). The rate of disproportionation by the sol–gel encapsulated ion pair was found to depend on the sizes and steric constraints of the starting materials. For instance, under comparable conditions the respective conversions of 1,3-cyclohexadiene, 1,2-dihydronaphthalene, 1,2-dihydro-4-methylnaphthalene, 1,2-dihydro-4-ethylnaphthalene, 1,2-dihydro-8-methoxynaphthalene, 1,2-dihydrophenanthrene, 3,4-dihydrophenanthrene, and 10,11-dihydrobenz[*a*]anthracene were 96, 99, 41, 17, 87, 73, 50, and 23%. This set of experiments emphasizes the difference between catalysis by zeolites and sol–gel catalysis. Zeolites are usually limited to a single substrate, whereas sol–gel catalysts can be utilized for a family of substrates, albeit with different efficiencies and rates. There are advantages and specific applications to each of these types of catalysts. The sol–gel entrapped $[(C_8H_{17})_3NMe]^+[RhCl_4 \cdot nH_2O]^-$ was found to be perfectly leach-proof during the disproportionation processes (amount of leached metal $\ll 1$ ppm) and could be recycled for 10–20 runs.

The second case for catalytic hydrogen transfer is double bond migration (Parshall and Ittel 1992). An example is 1-octen-3-ol which can be catalytically isomerized by sol–gel-entrapped $RhCl(PPh_3)_3$ to 3-octenone. Entrapment in a silica sol–gel matrix increases the catalytic activity. The catalytic properties of the encapsulated catalyst were found to depend on the structural features of the sol–gel matrix, and a rough correlation exists between the catalytic activity and the surface area. Thus, when the components of $[(C_8H_{17})_3NMe]^+[RhCl_4 \cdot nH_2O]^-$, ($RhCl_3 \cdot 3H_2O$ and $[(C_8H_{17})_3NMe]Cl$) have been entrapped in sol–gel matrices prepared from tetramethoxy-, tetraethoxy-, tetrapropoxy- and tetrabutoxysilane (TMOS, TEOS, TPOS, and TBOS, respectively), under identical conditions, the N_2-Langmuir surface areas of the corresponding immobilized catalysts were 722, 539, 975, and 603 m^2g^{-1} (the average pore diameter was 14–15 Å in the four matrices), and the respective initial rates of allylbenzene isomerization by these catalysts were found to be 1.00, 0.23, 2.10, and 1.40 $mmol\,min^{-1}\,ml^{-1} \times 10^{-2}$.

Catalytic hydrogenations and hydrogenolyses by sol–gel entrapped ion-pairs are further examples, and these are described in Blum *et al.* (1996). Since the sol–gel entrapped catalysts, e.g., $[(C_8H_{17})_3NMe]^+[RhCl_4 \cdot nH_2O]^-$, withstand temperatures as high as 130 °C and hydrogen pressures as high as 60 atm, conditions under which the catalyst decomposes in solution, one can hydrogenate unsaturated substrates that are not affected by the nonencapsulated catalyst, a typical example being benzonitrile (Blum *et al.* 1996). Of environmental importance are the dehalogenation reactions. Thus, fluoro- and chloroarenes undergo hydrogenolysis with the catalyst at 16.5 atm hydrogen to give halogen-free cyclohexane derivatives.

Sol–gel entrapped chiral catalysts that include (*R*)-(+)- and (*S*)-(−)-2,2′-[bis(diphenylphosphino)-1,1′-binaphthyl]chloro(*p*-cymene)ruthenium chloride (Ru-BINAP), [(4*S*,5*S*)-(+)- and [(4*R*,5*R*)-(−)-[4,5-bis(diphenylphosphinomethyl)-2,2′-dimethyl-1,3-dioxalene]chlororhodium (Rh-DIOP), and [(2*S*,4*S*)-(−)-1-*tert*-butoxycarbonyl-4-diph enylphosphino-2-(diphenylphosphinomethyl)pyrrolidine]chlororhodium, (Rh-BPPM) (Gelman *et al.* 1999) have been used for asymmetric hydrogenation of the prochiral itaconic acid (Kawano *et al.* 1987). Thus, enantioselective hydrogenation of itaconic acid with Rh-BPPM entrapped in silica sol–gel resulted in leach-proof material which catalyzed quantitatively itaconic acid into (−)-2-methylsuccinic acid of 78% optical purity, the value reported for the same reaction using the nonimmobilized Rh-BPPM catalyst (Ojima and Kogure 1978). The ability to employ the heterogenized complexes in an aqueous medium should be noticed here, since their nonentrapped, water-insoluble analogues are completely inactive in this environmentally friendly solvent.

The transparency of silica sol–gel material makes them particularly suitable for photocatalysis. Remaining in the organometallic catalysis domain, a representative case is the photocatalyzed reduction of protons to hydrogen, using ruthenium and iridium bipyridyl complexes coentrapped as an isolated donor–acceptor pair in a silica sol–gel matrix (Slama-Schwok *et al.* 1991). For a summary of representative sol–gel photocatalyses see Castelano and Meyer (1997).

3. Biochemical Reactivity with Entrapped Proteins

One of the fastest growing fields within doped sol–gel materials is the entrapment of purified enzymes, cell extracts, whole-cells, antibodies, and other proteins (Livage 1996). Various applications of these novel bioreactive materials have been developed including biosensors (see Sect. 4), immunoadsorbents, optical components, environmentally-related materials, and biotechnology-related materials. Some of the properties that have contributed to the attractiveness of this immobilization methodology are the enhanced stability of the encapsulated molecules (Zheng *et al.* 1998), the absolute prevention of leaching of the proteins due to the effective caging, the ability to immobilize without resorting to covalent modification of the protein, and protection against enzymatic autodigestion. One of the most studied enzymes in this context has been glucose oxidase, reviewed elsewhere (Avnir *et al.* 1994a) (see also Sect. 4). Many other enzymes have been studied; some representative examples which demonstrate the scope and potential of this field are given below. (For additional examples see Avnir and Braun 1996, Dunn *et al.* 1998.)

3.1 Alkaline Phosphatase

This was an early case of enzyme entrapment by the direct use at TMDS (Braun *et al.* 1990) which triggered many enzyme entrapments to come. When entrapped in SiO_2 sol–gel and dried, its dose–response curve as determined by the phosphate bond cleavage of *p*-nitrophenyl phosphate did not fit a Michaelis–Menten model, but reflected a distribution of activity efficiencies of the individually trapped molecules. These variations in activity may be the result of different accessibilities of the individual active sites to the pore network in which the substrate resides, or the result of various degrees of denaturation caused by the sol–gel process. However, it was later found in many studies that when the gel was kept wet, the linear Michaelis–Menten behavior was retained. In any event, even in its dried entrapped form, alkaline phosphatase was proven to gain stability, and it retained its full activity for at least two months. One may attribute this enhanced stability to the protective nature of the cage itself and to the rigidity of the SiO_2 matrix which reduces the freedom of peptide-chain refolding molecular motions which occur in denaturation processes. For other representative studies of this group, particularly on trypsin see, e.g., Shtelzer *et al.* (1992).

3.2 Cytochrome *c*

Cytochrome *c* entrapped in silica sol–gel materials was studied in detail by Dunn *et al.* (Miller *et al.* 1996). The effects of several key parameters was studied (pH, alcohol/alkoxide ratio, buffer type) and it was found that the behavior of the protein encapsulated in silica gels mirrored their behavior in solution. This suggests that the stability limiting factor for encapsulating these biological molecules is determined by the solution stability of the protein. Entrapped cytochrome *c* was capable of undergoing redox processes exhibiting reversible changes in the oxidation state of iron upon addition of either hydrogen peroxide or sodium dithionate. Again, an important advantage of the sol–gel methodology was revealed: the thermal stability of cytochrome *c* was significantly improved by immobilization in a porous silica network (Lan *et al.* 1999).

3.3 Invertase

Invertase entrapment was used to demonstrate the usefulness of sol–gel materials beyond silica (Kurokawa 1996). Thus, gel fibers were formed from cellulose acetate and titanium as well as zirconium alkoxides. These proved to be stable in common solvents, phosphate solution, and electrolyte solution over a wide range of pH, from 4 to 10. The immobilized enzyme was found to be stable for long periods and was used in a packed column. The pH maximum of activity became more alkaline compared with the native enzyme, reflecting the amphoteric property of the metal oxide.

3.4 Calcium-binding Protein Oncomodulin

This was studied by Brennan and co-workers, providing a methodology for improving further the thermal stability of entrapped proteins (Zheng *et al.* 1998). These authors showed that the performance of the sol–gel-entrapped protein can be improved substantially by maximizing the protein's thermodynamic stability so that it can withstand the harsh conditions associated with entrapment. Two mutants of the protein were entrapped, and the thermodynamic stability of both proteins could be manipulated by adjusting the level of calcium ions present during entrapment.

3.5 Lipase

Lipase entrapment, studied by Reetz *et al.*, demonstrated that sol–gel matrices can greatly enhance the activity of the sol–gel immobilized enzyme, as compared with its solution activity (Reetz *et al.* 1996). The enhancement of activity was achieved by entrapping lipases in hydrophobic organic/inorganic hybrid silica gels, prepared by hydrolysis of alkyl-substituted silanes in the presence of the enzyme. The method is applicable for a wide variety of lipases, yielding immobilized lipases with esterification activities enhanced by a factor of up to 88 compared with the commercial enzyme powders under identical conditions. Studies of the stability of sol–gel immobilized lipases under reaction conditions or storage (dry or in aqueous or organic medium) revealed an excellent retention of enzymatic activity. "Magnetic lipase" was prepared by mixing lipose with a paste of ferrite powder, tetramethoxysilane, 3-aminopropyltriethoxysilane, and propyltrimethoxysilane (Kuncová and Sivel 1997).

3.6 Bacteriorhodopsin

Bacteriorhodopsin was used to demonstrate the advantages of entrapping a photoreactive protein (Weetall 1996). From it, purple membrane fragments were isolated and entrapped in dried sol–gel glass. Samples showing constant activity when stored at 4 °C lost further activity at room temperature, eventually reaching a level where the activity remained constant.

3.7 Alcohol Dehydrogenase

Alcohol dehydrogenase entrapped in silica sol–gel was shown to withstand harsh nonaqueous and gas-phase environments (Williams and Hupp 1998). Furthermore, it was demonstrated that cofactors can be used with entrapped enzyme—specifically, the pairs ADH/NAD$^+$ or ADH/NADH. The utilization of changes in fluorescence from the soluble, reduced cofactor NADH upon exposure to alcohols or aldehydes allows for semiquantitative determination of both substrates.

3.8 General Enzyme Entrapment

A general enzyme entrapment methodology has been described by Gill and Ballesteros (1998). The paper describes a class of sol–gel precursors based on polyol silicates and polyol siloxanes. In particular, poly(glyceryl silicate) provided high biocompatibility and mild encapsulation conditions, and enabled the reproducible and efficient confinement of a wide variety of proteins and cells inside silica. The methodology was extended to metallosilicate, alkylsiloxane, functionalized siloxane, and composite sol–gels, thereby allowing the fabrication of a physicochemically diverse range of bio-doped polymers. The hybrid materials displayed activities approaching those of the free enzymes and proteins, together with the high stabilities and robustness that characterize sol–gel bioceramics.

3.9 Cells and Whole Cell Extracts

These have been entrapped successfully in sol–gel matrices. An example is the silica sol–gel entrapment of parathion hydrolase extracted from *Pseudomonas* spp. (Dozorets *et al.* 1996). Another example is the study of Kaufmann and Mandelbaum (1998) in which atrazine chlorohydrolase from cell extracts of *Pseudomonas* was immobilized successfully in hydrophobic sol–gel matrices, and used for detoxification of this hazardous environmental pollutant. Carturan *et al.* (Inama *et al.* 1993) studied the sol–gel encapsulation of whole yeast cells (*Saccharomyces cerevisiae* as spores) using SiO_2 sols. This prehydrolysis approach, also used for enzyme entrapment, e.g., by Johnson (Johnson and Whateley 1971), predated the advent of the sol–gel metal alkoxide polymerization procedures.

3.10 Antibodies

Antibodies entrapped in sol–gel materials comprise our last but very interesting case (Aharonson *et al.* 1994). An application as an immunochromatographic material was demonstrated by Knopp *et al.* (Fröschl *et al.* 1998) who constructed a polyclonal-based 1-nitropyrene enzyme-linked immunosorbent assay (ELISA) methodology for environmental analytical chemistry for screening purposes. A different configuration was demonstrated by Roux *et al.* (1997). These authors succeeded in the encapsulation of antigens, and showed that the pores of the sol–gel matrix are large enough to allow the diffusion of antibodies through the gel. They successfully performed ELISA immunoassays using hydatid cyst fluid as the source of antigens and sera from human patients as the source of antibodies. Finally, Shabat *et al.* (1997) demonstrated the first successful entrapment of catalytic antibodies. These materials, which in this case were used for the continuous-flow column synthesis of optically pure enantiomers, carry great potential for many other syntheses.

4. Sol–Gel Sensors

A chemical sensor is a small device comprising a molecular recognition system embedded within or on a supporting matrix and equipped with a signal translation and quantitation capability. Extensive sol–gel research efforts have been devoted to the optimization of each of these elements for different end applications and for improving the synergism between the inorganic and organic ingredients of the hybrid sensors (Collinson 1998). Sol–gel doping provides opportunities to benefit from the very large number of organic and biochemical molecular recognition reagents, including high selectivity, and still benefit from the favorable physical properties of silicates and Ormosils (organically modified silicas) as detailed earlier.

4.1 Molecular Recognition

The first series of reports on sol–gel sensing demonstrated the ability to incorporate molecular recognition elements in sol–gel-derived silicate monoliths by sol–gel doping (Zusman *et al.* 1990, Braun *et al.* 1990, Ellerby *et al.* 1992, Rottman *et al.* 1992, Lev 1992, Iosefzon *et al.* 1992, Dave *et al.* 1994). Colorimetric reagents were encapsulated in sol–gel-derived silicates, and color evolution was monitored after exposure of the doped monolith to liquid or gaseous environments. These reports illuminated two fundamental properties of doped sol–gel materials, which make them suitable candidates for sensing applications. They allow penetration of foreign molecules from the surrounding gas or liquid phase into the gel and further enable the participation of the diffusing compounds in molecular recognition interactions within the glass network. Proton transfer reactions (e.g., with pH indicators), reversible or irreversible redox reactions, complex formation, and enzymatic reactions involving diffusing analytes and organic dopants were demonstrated in the porous solid phase. The first reports on sol–gel sensing also made an important distinction between continuous and disposable sensors, which can both be realized by sol–gel doping technology (Lev 1992). Later, successful entrapment of immunoglobulins in sol–gel matrices opened the door for optical and electrochemical immunoassay—ELISA-type—detection (Roux *et al.* 1997, Turniansky *et al.* 1996, Fröschl *et al.* 1998, Wang *et al.* 1998).

It was soon found that the porous sol–gel matrix provides sufficient connectivity to enable communication between different functionalities that are not necessarily embedded in the same pore. Multiligand chelate formation and multistep reactions within the glass phase exemplify this property. For example, iron(II) forms red, three-ligand chelates with 1,10-phenanthroline. This is a convenient analytical method for the determination of iron in aqueous environments. Electron spectroscopy shows that the very same chelate is formed within the glass phase when 1,10-phenanthroline-doped sol–gel glasses are immersed in aqueous solution containing ferrous cations (Iosefzon-Kuyavskaya *et al.* 1992). The ability to carry out multiple reactions employing different catalysts or biocatalysts within the sol–gel matrix is another attractive feature. This is used in integrated sol–gel sensors, where different dopants participate or catalyze different reactions (Lev *et al.* 1992, Wolfbeis *et al.* 1996, Avnir 1995). For example, optical (Avnir *et al.* 1994b) and amperometric (CocheGuerente *et al.* 1997) glucose sensing can be carried out by sol–gel materials doped with glucose oxidase, horseradish peroxidase, and a chromophore or charge mediator. Glucose is oxidized by the flavin prosthetic group of the glucose oxidase, which is then regenerated by oxygen to give hydrogen peroxide. The hydrogen peroxide is further biocatalytically reduced by a dye-forming reaction or by a charge mediator that shuttles the electrons to a conductive part of the electrode. Multienzyme catalysis is now routinely used for sensing applications (Avnir *et al.* 1994a).

4.2 Tailoring Sol–gel Composition for Specific End Applications

It was soon found that specific preparation protocols and different composition of the Ormosils should be tailored in order to meet the specifications of target end applications. This can be exemplified by the innovative remedies devised for a major problem that was raised in sol–gel infancy—the leaching of the dopants from sol–gel sensors during prolonged operation. Altering the dopant or the supporting matrix, and the interaction between the two solved this problem in many cases. It is relatively simple to construct leak-free sol–gel films and monoliths for high molecular weight proteins and synthetic polymer dopants. This approach was used to construct an electrochemical sensor for glucose using glucose oxidase-bound ferrocene (Sampath *et al.* 1996a). Osmium redox polymer was also embedded in sol–gelderived electrochemical biosensors for hydrogen peroxide (Park *et al.* 1996). Zero leaching of the ferrocene and osmium mediators was reported even after prolonged operation periods.

Altering the silicate matrix was another successful approach. Carrying out the sol–gel polymerization under low pH gives rise to extremely small pore-size gels, which prevent leakage of the pH indicators (Aharonson *et al.* 1994). This approach is successful only for low-molecular weight analytes. McDonagh *et al.* (1998a) devised an innovative solution to prevent leakage of ruthenium complex dyes from their oxygen sensors. Quenching of the ruthenium fluorescence by oxygen was used to quantify the dissolved oxygen level. The ruthenium complex was embedded in hydrophobic methyltriethoxysilane-derived matrix. The hydrophobic matrix enhanced penetration of the uncharged oxygen analyte, whereas water was excluded from the sensing element.

Another approach to prevent leaching is to increase the affinity between the dopant and the silicate network. This can be accomplished by designing Ormosil networks containing functional groups that bind the dopant by electrostatic or hydrophobic interactions (e.g., Kraus *et al.* 1993, Lobnik and Wolfbeis 1998). Covalent bonding of reagents to pre-prepared sol–gel materials or to silyl monomers is increasingly used to prevent leaching from optical (Aharonson *et al.* 1994) and electrochemical sensors (Gun and Lev 1996).

4.3 Optical Sensing

Optical sol–gel sensing has matured since the first colorimetric feasibility studies were demonstrated in the early 1990s. Practically all major types of photometric sensing, including absorbance, plasmon resonance, fluorescence, room-temperature phosphorescence, and chemiluminescence spectroscopies, have been successfully carried out in doped sol–gel silicates (e.g., Wolfbies *et al.* 1996, Avnir 1995, Lev *et al.* 1995). Several examples depicted in Fig. 1 demonstrate the versatility in tailoring optical sensors of various configurations.

(*a*) *Fiber optic sensors*. Three modes of fiber optic photometric sensors have been proposed. In the simplest mode, optical fibers are used only to guide the light from the doped sol–gel sensor (Fig. 1(a)). This approach was used for pH sensing (Rottman *et al.* 1992). Grattan *et al.* (1991) used a similar approach for fluorescence pH detection with fluorescein isothiocyanate indicator (Fig. 1(b)). A related technique based on the modification of the optical fiber itself has also been developed. An unclad section of a home-made, porous optical fiber was infiltrated with the sol–gel precursors and pH indicators. This configuration (Fig. 1(c)) benefits from a long optical pathway, short analyte diffusion length, and simple

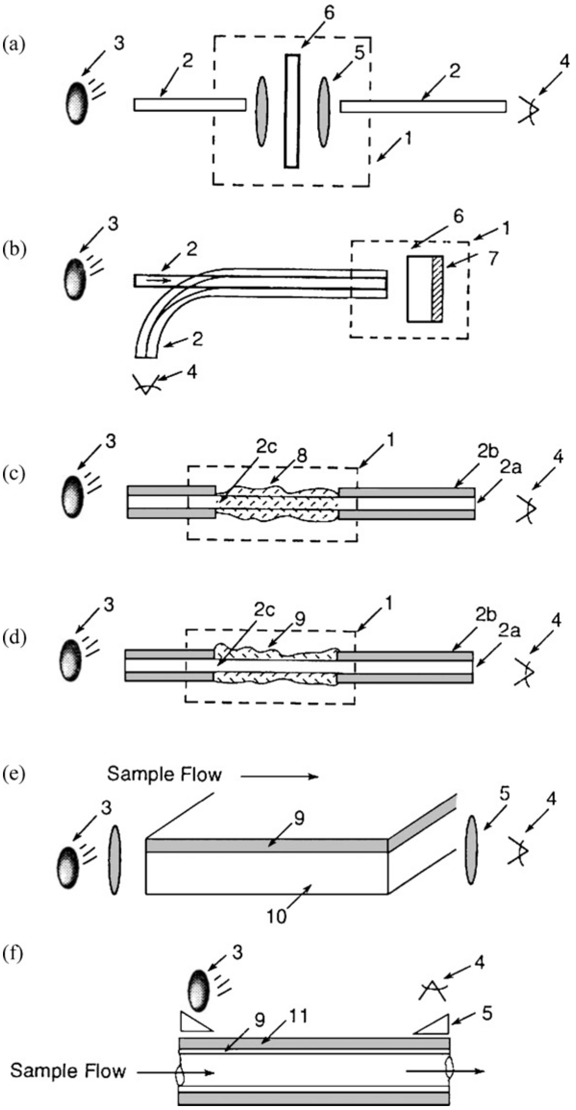


Figure 1
Configuration of reagent-doped sol–gel waveguides (after Lev *et al.* 1995, with permission). (a) Direct mode transmission sensor (Rottman *et al.* 1992); (b) fluorescence- or reflection-mode sensor (Grattan *et al.* 1991); (c) sol–gel-modified porous optical sensor (Shariari *et al.* 1994); (d) evanescent optical fiber sensor (McDonagh *et al.* 1998a); (e) planar waveguide evanescent sensor (Yang *et al.* 1994); (f) capillary flow-through evanescent sensor (Weigl and Wolfbeis 1994). 1: Sample; 2: optical fiber; 2a: core; 2b: cladding; 2c: uncladded section; 3: source; 4: photodetector, 5: auxiliary coupling; 6: sol–gel sensor; 7: optical mirror; 8: doped sol–gel glass; 9: doped sol–gel film; 10: planar glass support; 11: glass capillary.

optical coupling auxiliaries. A third approach, developed by McDonagh *et al*. (1998a, 1998b) is based on evanescent wave fluorescence excitation and involves pH sensitive sensors made by dip-coating unclad, fused-silica optical fiber with a fluorescein-doped sol–gel film (Fig. 1(d)). Such evanescent wave sensors benefit from a long optical path length, relatively low dye bleaching in laser-irradiated fiber optic sensors, and the large contact area between the film and the examined environment.

(*b*) *Other waveguide configurations*. Sensing applications often require irregular sensor configurations; for example, planar and capillary tube waveguide configurations are useful in flow-cell detectors. Yang *et al*. (1994) devised a planar waveguide pH sensor made from a microscope slide coated with a bromophenol blue-doped thin silica film (Fig. 1(e)). In this configuration the microscope slide operates as a waveguide and the doped film as a sensing cladding. In another application Weigl and Wolfbeis (1994) proposed a sol–gel-coated capillary sensor in which the capillary tube acts as a core, as a reagent-doped coating, and as a cladding (Fig. 1(f)). This configuration can be used for flow-through detectors and also as disposable capillary suction detectors that provide accurate sampling volumes.

(*c*) *Near-field optical sensors*. A demonstration of a way to push miniaturization of doped sol–gel sensors to the limit is the near-field optical sensor. Samuel *et al*. (1994) polymerized a pyranine-doped silicate microdot on the tip of a drawn glass capillary. A light beam was directed through the capillary with the sensing tip inducing pH-dependent fluorescence that was detected by a stationary photodetector. This device exhibited a response time of less than one second and reversible pH detection. Near-field microscopy opens the door to intracellular sensing and micron-resolution spatial pH mapping.

4.4 *Electrochemical Sensing*

The rapid development of optical doped sol–gel applications was soon followed by a parallel development of electroanalytical applications. Several reviews discuss the progress in this field (Lev *et al*. 1997, Walcarius 1998, Tess and Cox 1999, Collinson 1998, Alber and Cox 1997). Despite the fact that the diffusion coefficients of organic compounds in the silicate films are very low, the feasibility of sol–gel electrochemistry and bioelectrochemistry with doped silicates was soon proven. Dvorak and Armond (1993) made use of the favorable optical transparency of silicate films to demonstrate the photoelectrochemical activity of tris(2,2′-bipyridyl)ruthenium. The observed quantum yield was low and the electrochemical response was rather sluggish but feasibility was indeed demonstrated. Tatsu *et al*. (1992) constructed the first sol–gel glucose flow-injection analyzer using glucose oxidase-doped inorganic silicate powder which was attached to an oxygen electrode by an external cellulose membrane. Oxygen depletion was used for quantification of the converted glucose.

Since these initial feasibility studies, electroanalytical studies have shifted towards doped Ormosils and composite materials which seem to provide higher diffusion rates and less sluggish electrochemistry. Indeed, within a relatively short period of time several potentiometric (e.g., Kim *et al*. 1997, Kimura *et al*. 1996), amperometric (e.g., Hsueh and Collinson 1997, Hu *et al*. 1998), and bioelectrochemical (e.g., Audebert *et al*. 1993, Narang *et al*. 1994, Pandey *et al*. 1999) applications where demonstrated for sol–gel-derived electrodes. Moreover, several reports of permselectivity of native and doped silicate and Ormosil modified electrodes were published (e.g., Collinson 1998).

4.5 *Composite Materials Employing More Than One Phase*

Increasing attention is also being devoted to composite materials employing more than one phase. For example, carbon ceramic composite electrodes (CCEs) and related sol–gel composites are attracting increased interest due to their versatility and improved electrochemical performance (e.g., Tsionsky *et al*. 1994, Gun *et al*. 1994a, 1994b, Wang *et al*. 1996). The basic ingredients of CCEs are graphite or carbon black powders dispersed in an Ormosil (e.g., methylsilicate) sol–gel network. The carbon or graphite powder provides electrical conductivity by an electron percolation mechanism, and the porous silicate matrix contributes rigid and brittle construction. The hydrophobicity of the methyl modifiers prevents water penetration into the electrode, and thus only the outermost section of the electrode remains in contact with the electrolyte. These electrodes showed an improvement of up to three orders of magnitude in Faradaic signal/background (capacitive) noise ratios compared with glassy carbon electrodes (Gun *et al*. 1994a, 1994b). The electrodes can be molded in diverse configurations, including supported or unsupported thick films, rods, disks, and even in the form of microelectrodes ($\sim 20\,\mu\text{m}$ in diameter). Graphite exhibits poor electrocatalytic activity, but addition of trace-metal or organometallic catalysts improves this activity. CCEs containing organometallic catalysts, such as cobalt phthalocyanine or cobalt porphyrin, showed pronounced electrocatalytic activity toward the reduction of dioxygen (Tsionsky and Lev 1995a, 1995b, Wang *et al*. 1997a). Since the electrodes are porous, the inert metal and organometallic-modified CCEs can act as electrodes for gases such as oxygen, carbon monoxide, and sulfur dioxide (Tsionsky *et al*. 1995a).

Since CCEs are molded at room temperature, it is possible to incorporate enzymes within the silicate material (Pankratov and Lev 1995, Tsionsky *et al.* 1994, Wang *et al.* 1996). The simplest form of enzyme electrode comprises a glucose oxidase–methyl silicate (or silica)–graphite composite. Dissolved oxygen serves as an electron acceptor, and the resulting hydrogen peroxide is electro-oxidized on the electrode. In order to account for the over-voltage for hydrogen peroxide oxidation, a palladium catalyst was added to the above construction. Electrodes for lactate (with lactate oxidase) and amino acids (amino acid oxidase) were prepared in a similar manner (Sampath and Lev 1996b). Reduction of the signal dependence on the level of oxygen was achieved by incorporating a mediator into the sol–gel starting solution. Tetrathiafulvalene and glucose oxidase-doped CCE was found to show an oxygen-independent response (Pankratov and Lev 1995). In other modifications the ferrocenyl groups were covalently bound to the glucose oxidase enzyme, to the carbon powder, or bonded to the sol–gel precursor (Gun *et al.* 1995). The CCE biosensors were then molded by the usual protocol. In all cases zero leaching and good stability were observed. Wang *et al.* (1996) showed that CCEs can be produced by thick-film ink-jet technology. This technology enables highly reproducible mass production of biosensors. Wang and Pamidi (1997b) have also successfully replaced the graphite powder by noble metal particles, and our group has devised ways to replace the graphite by noble metal nanoparticles that are stabilized by aminosilane capping shells (Bharathi and Lev 1998, 1999).

Bibliography

Aharonson N, Altstein M, Avidan G, Avnir D, Bronshtein A, Lewis A, Liberman M, Ottolenghi M, Polevaya Y, Rottman C, Samuel J, Shalom S, Strinkovski A, Turniansky A 1994 Recent developments in organically doped sol–gel sensors: a microns-scale probe; successful trapping of purified polyclonal antibodies; solutions to the dopant-leaching problem. *Mater. Res. Soc. Symp. Proc.* **346**, 519–30

Alber K S, Cox J A 1997 Electrochemistry in solids prepared by sol–gel processes. *Mikrochim. Acta* **127**, 131–47

Audebert P, Demaille C, Sanchez C 1993 Electrochemical probing of the activity of glucose-oxidase embedded sol–gel matrices. *Chem. Mater.* **5**, 911–3

Avnir D 1995 Organic chemistry within ceramic matrices: doped sol–gel materials. *Acc. Chem. Res.* **28**, 328–34

Avnir D, Braun S 1996 *Biochemical Aspects of Sol–gel Science and Technology*. Kluwer, Boston

Avnir D, Braun S, Lev O, Levy D, Ottolenghi M 1994b Organically-doped sol–gel glasses: optics, photophysics and chemical sensing. In: Klein L C (ed.) *Sol–Gel Optics—Processing and Applications*. Kluwer, Boston, Chap. 23, pp. 539–82

Avnir D, Braun S, Lev O, Ottolenghi M 1994a Enzymes and other proteins entrapped in sol gel materials. *Chem. Mater.* **6**, 1605–14

Avnir D, Klein L C, Levy D, Schubert U, Wojcik A B 1998 Organo-silica sol–gel materials. In: Apeloig Y, Rappoport Z (eds.) *The Chemistry of Organosilicon Compounds*. Wiley, Chichester, UK, Vol. 2, Chap. 40, PP. 2317–62

Bharathi S, Lev O 1998 Sol–gel-derived nanocrystalline gold-silicate composite biosensor. *Anal. Commun.* **34**, 29–31

Bharathi S, Lev O 1999 Electrodeposition of thin gold films from an aminosilicate stabilized gold sol. *Electrochem. Solid State Lett.* **2**, 284–7

Bianchini D, Lacross M, Pajonk J M, Teichner J 1981 Spilled-over hydrogen transport from platinum-on-alumina catalyst to methoxylated silica aerogel. *J. Catal.* **64**, 411–18

Blum J, Rosenfeld A, Polak N, Israelson O, Schumann N, Avnir D 1996 Comparison between homogeneous and sol–gel-encapsulated rhodium-quaternary ammonium ion pair catalysts. *J. Mol. Catal. A: Chem.* **107**, 217–23

Braun S, Rappoport S, Zusman R, Avnir D, Ottolenghi M 1990 Biochemically active sol–gel glasses: the trapping of enzymes. *Mater. Lett.* **10**, 1–5

Brinker C J, Scherer G W 1990 *Sol–Gel Science*. Academic Press, New York

Castelano F N, Meyer G J 1997 Light induced processes in molecular gel materials. *Prog. Inorg. Chem.* **44**, 167–208

Coche-Guerente L, Cosnier S, Labbe L 1997 Sol–gel derived composite materials for the construction of oxidase/peroxidase mediatorless biosensors. *Chem. Mater.* **9**, 1348–52

Collinson M M 1998 Analytical applications of organically modified silicate. *Mikrochim. Acta* **129**, 149–65

Dave B C, Dunn B, Valentine J S, Zink J I 1994 Sol–gel encapsulation methods for biosensors. *Anal. Chem.* **66**, 1120-117A

Dickey F H 1949 The preparation of specific adsorbents. *Proc. Natl. Acad. Sci.* **35**, 227–37

Dosoretz C, Armon R, Starovetsky J, Rothchild N 1996 Entrapment of parathion hydrolase from Pseudomonas spp. in sol–gel glass. *J. Sol–Gel Sci. Technol.* **7**, 7–12

Dvorak O, De Armond M K 1993 Electrode modification by the sol–gel method. *J. Phys. Chem.* **97**, 2646–8

Dunn B, Miller J M, Dave B C, Valentine J S, Zink J I 1998 Strategies for encapsulating biomolecules in sol–gel matrices. *Acta Mater.* **46**, 737–41

Ellerby L M, Nishida C R, Nishida F, Yamanaka S A, Dunn B, Valentine J S, Zink J I 1992 Encapsulation of proteins in transparent porous silicate-glasses prepared by the sol–gel method. *Science* **225**, 1113–6

Fröschl B, Knopp D, Niessner R 1998 Possibilities and limitations of an enzyme-linked immunosorbent assay for the detection of 1-nitropyrene in air particulate matter. *Fresenius J. Anal. Chem.* **360**, 689–92

Gelman F, Avnir D, Schumann H, Blum J 1999 Sol–gel entrapped chiral rhodium and ruthenium complexes as recyclable catalysts for the hydrogenation of itaconic acid. *J. Mol. Catal. A: Chem.* **146**, 123–5

Gill I, Ballesteros A 1998 Encapsulation of biologicals within silicate, siloxane, and hybrid sol–gel polymers: an efficient and generic approach. *J. Am. Chem. Soc.* **120**, 8587–98

Gonzales R D, Lopez T, Gomez R 1997 Sol–gel preparation of supported metal catalysts. *Catal. Today* **35**, 293–317

Grattan K T V, Badini G E, Palmer A W, Tseung A C C 1991 Use of sol–gel techniques for fiberoptic sensor applications. *Sensors Actuators A* **25–7**, 483–7

Gun G, Tsionsky M, Lev O 1994b Voltammetric studies of carbon-silica composite working electrodes. *Anal. Chim. Acta* **294**, 261–70

Gun G, Tsionsky M, Lev O 1994a Preparation and characterization of carbon ceramic electrodes. *Mater. Res. Soc. Symp. Proc.* **346**, 1011–6

Gun J, Lev O 1996 Sol–gel derived, ferrocenyl-modified silicate—graphite composite electrode: wiring of glucose oxidase. *Anal. Chim. Acta* **336**, 95–106

Gun J, Tsionsky M, Rabinovich L, Golan Y, Rubinstein I, Lev O 1995 Electrochemical and morphological characterization of palladium modified carbon-ceramic electrodes. *J. Electroanal. Chem.* **395**, 57–66

Han L, Niemczyk T M, Haaland D M, Lopez G P 1999 Enhancing IR detection limits for trace polar organics in aqueous solutions with surface-modified sol–gel-coated ATR sensors. *Appl. Spectrosc.* **53**, 381–9

Hsueh C C, Collinson M M 1997 Permselectivities and ion-exchange properties of organically modified sol–gel electrodes. *J. Electroanal. Chem.* **420**, 243–9

Hu Z M, Seliskar C J, Heineman W R 1998 Voltammetry of $[Re(DMPE)_3]^+$ at ionomer-entrapped composite-modified electrodes. *Anal. Chem.* **70**, 5230–6

Inama L, Dire S, Carturan G, Carazza A 1993 Entrapment of viable microorganisms by SiO_2 sol–gel layers. *J. Biotechnol.* **30**, 197–210

Iosefzon-Kuyavskaya B, Gigozin I, Ottolenghi M, Avnir D, Lev O 1992 Spectrophotometric detection of heavy metals by doped sol–gel glass detectors. *J. Non-Cryst. Solids* **147/8**, 808–12

Johnson P, Whateley T L 1971 On the use of polymerizing silica gel systems for the immobilization of trypsin. *J. Coll. Interface Sci.* **37**, 557–63

Jorgensen C K, Reisfeld R 1996 *Optical and Electronic Phenomena in Sol–Gel Glasses and Modern Applications*. Springer, Berlin

Kauffmann C, Mandelbaum R T 1998 Entrapment of atrazine chlorohydrolase in sol–gel glass matrix. *J. Biotechnol.* **62**, 169–76

Kawano W, Ishii Y, Ikariya T, Saburi M, Yoshikawa S, Uchida Y, Kumobayashi H 1987 Ruthenium(II)-BINAP complex catalyzed asymmetric hydrogenation of unsaturated dicarboxylic acids. *Tetrahedron Lett.* **28**, 1905–8

Kim W, Chung S, Park S B, Lee S C, Kim C, Sung S D 1997 Sol–gel method for the matrix of chloride-selective membranes doped with tridodecylmethylammonium chloride. *Anal. Chem.* **69**, 95–8

Kimura K, Sunagawa T, Yokoyama M 1996 Chemical modification of sol–gel-derived glass by a neutral carrier for ion sensors. *Chem. Commun*, 745–6

Klein L C (ed.) 1994a *Sol–Gel Optics—Processing and Applications*. Kluwer, Boston

Kraus S C, Czolk R, Reichert J, Ache A J 1993 Optimization of the sol–gel process for the development of optochemical sensors and actuators. *Sensors Actuators B* **15/16**, 199–202

Kuncová G , Sivel M 1997 Lipase immobilized in organic-inorganic matrices. *J. Sol–Gel Sci. Technol.* **8**, 667–71

Kurokawa Y 1996 Entrap-immobilization of enzyme on composite gel fiber using a gel formation of cellulose acetate and metal (Ti, Zr) alkoxide. *Polym. Gels Networks* **4**, 153–63

Labroue D, Poilblanc R 1977 Homogeneous hydrogenation of styrene with metal carbonyl clusters. *J. Mol. Catal.* **2**, 329–33

Lan E H, Dave B C, Fukuto J M, Dunn B, Zink J I, Valentine J S 1999 Synthesis of sol–gel encapsulated heme proteins with chemical sensing properties. *J. Mater. Chem.* **9**, 45–53

Lev O 1992 Diagnostic applications of sol–gel glasses doped with organic reagents. *Analysis* **20**, 543–54

Lev O, Iosefson-Kuyavskaya B, Gigozin I, Ottolenghi M, Avnir D 1992 Glass phase photometry, high sensitivity spectrophotometric method: 1. Determination of dissolved iron. *Fresenius J. Anal. Chem.* **343**, 370–2

Lev O, Tsionsky M, Rabinovich L, Glezer V, Sampath S, Pankratov I, Gun G 1995 Organically modified sol–gel sensors. *Anal. Chem.* **67**, 22A–31A

Lev O, Wu Z, Bharathi S, Glezer V, Modestov A, Gun J, Rabinovich L, Sampath S 1997 Sol–gel materials in electrochemistry. *Chem. Mater.* **9**, 2354–75

Levy D, Del Monte F, Quintana X, Oton J M 1997 Color displays with gel glass dispersed liquid crystals. *J. Sol–Gel Sci. Technol.* **8**, 1063–6

Levy D, Einhorn S, Avnir D 1989 Applications of the sol/gel process for the preparation of photochromic information recording materials: synthesis, properties, mechanisms. *J. Non-Cryst. Solids* **113**, 137–45

Livage J 1996 Bioactivity in sol–gel glasses. *C.R. Acad. Sci. Paris* **322**, 417–27

Livage J 1998 Sol–gel synthesis of heterogeneous catalysts from aqueous solutions. *Catal. Today* **41**, 3–19

Lobnik A, Wolfbeis O S 1998 Sol–gel based optical sensor for dissolved ammonia. *Sensors and actuators B Chem.* **51**, 203–7

Lorenz A, Kickelblick D, Schubert U 1997 Metal complexes in inorganic matrixes, 18. Phosphanyl-substituted titanium and zirconium alkoxides for tethering metal complexes in titania or zirconia and the x-ray structure analysis of polymeric $Zr(OPr)(O_3SiMe)_3$. *Chem. Mater.* **9**, 2551–60

Makote R, Collinson M M 1998 Template recognition in inorganic-organic hybrid films prepared by the sol–gel process. *Chem. Mater.* **10**, 2440–5

McDonagh C, MacCraith B D, McEvoy A K 1998b Tailoring of sol–gel films for optical sensing of oxygen in gas and aqueous phase. *Anal. Chem.* **70**, 45–50

McDonagh C M, Shields A M, McEvoy A K, MacCraith B D, Gouin J F 1998a Optical sol–gel-based dissolved oxygen sensor: progress towards a commercial instrument. *Sol–Gel Sci. Technol.* **13**, 207–11

Miller J M, Dunn B, Valentine J S, Zink J I 1996 Synthesis conditions for encapsulating cytochrome *c* and catalase in SiO_2 sol–gel materials. *J. Non-Cryst. Solids* **202**, 279–89

Monaco S J, Ko E I 1998 The puzzle of sol–gel catalysts. *CHEMTECH* **28** (6), 23–30

Narang U, Prasad P N, Bright F V, Ramanathan K, Kumar N D, Malhotra B D, Kamalasanan M N, Chandra S 1994 Glucose biosensor based on a sol–gel-derived platform. *Anal. Chem.* **66**, 3139–44

Negishi N, Fujino M, Yamashita H, Fox M A, Anpo M 1994 Photophysical and photochemical stability of rhodamine B encapsulated in SiO_2 matrices by the sol–gel method. *Langmuir* **10**, 1772–6

Ojima L, Kogure T 1978 Effective asymmetric hydrogenation of itaconic acid catalyzed by neutral and cationic rhodium(I) complexes with chiral pyrrolidino-diphosphine ligand. *Chem. Lett.* 567–8

Pandey P C, Upadhyay S, Pathak H C 1999 A new glucose biosensor based on sandwich configuration of organically modified sol–gel glass. *Electroanal.* **11**, 59–64

Pankratov I, Lev O 1995 Sol gel derived renewable-surface bio. *J. Electroanal. Chem.* **393**, 35–41

Park T M, Iwuoha I, Smyth M R, MacCraith B D 1996 Sol–gel-based amperometric glucose biosensor incorporating an osmium redox polymer as mediator. *Anal. Commun.* **33**, 271–3

Parshall G W, Ittel S D 1992 *Homogeneous Catalysis*. Wiley, New York pp. 9–24

Petit-Dominguez M D, Shen H, Heineman W R, Seliskar C J 1997 Electrochemical behavior of graphite electrodes modified by spin-coating with sol–gel entrapped ionomers. *Anal. Chem.* **69**, 703–10

Reetz M T, Zonta A, Simpelkamp J, Könen W 1996 *In situ* fixation of lipase-containing hydrophobic sol–gel materials on sintered glass—highly efficient heterogeneous biocatalysts. *Chem. Commun.* 1397–8

Rosenfeld A, Blum J, Avnir D 1996 Disproportionation of dihydroarenes by sol–gel entrapped $RhCl_3$-quaternary ammonium ion pair catalysts. *J. Catal.* **164**, 363–8

Rottman C, Ottolenghi M, Zusman R, Lev O, Smith M, Gong G, Kagan M L, Avnir D 1992 Doped sol–gel glasses as pH sensors. *Mater. Lett.* **13**, 293–8

Roux C, Livage J, Farhati K, Monjour L 1997 Antibody-antigen reactions in porous sol–gel matrices. *J. Sol–Gel Sci. Technol.* **8**, 663–6

Sampath S, Lev O 1996a Renewable, reagentless glucose sensor based on a redox modified enzyme and carbon silica composite. *Electroanalysis* **8**, 1112–6

Sampath S, Lev O 1996b Inert metal modified composite ceramic- carbon amperometric biosensors: Renewable controlled reactive layer. *Anal. Chem.* **68**, 2015–21

Sampath S, Lev O 1997 3-D organized self assembled electrodes: Enhanced communication between enzymes and electrodes. *Adv. Mater.* **9**, 410–3

Samuel J, Strinkovski A, Shalom S, Lieberman K, Ottolenghi M, Avnir D, Lewis A 1994 A metallized micropipette tip filled with doped xerogel for near field optical pH sensing. *Mater. Lett.* **21**, 431–4

Schumann H, Hasan M, Gelman F, Avnir D, Blum J 1998 Synthesis, immobilization and catalytic activity of some silylated cyclopentadienyl rhodium(I) complexes. *Inorg. Chim. Acta* **280**, 21–5

Sertchook H, Avnir D, Blum J, Joó F, Kathó A, Schumann H, Weimann R, Wernik S 1996 Sol–gel entrapped lipophilic and hydrophilic ruthenium-, rhodium-, and iridium-phosphine complexes as recyclable isomerization catalysts. *J. Mol. Catal. A: Chem.* **108**, 153–60

Shabat D, Grynszpan F, Saphier S, Turniansky A, Avnir D, Keinan E 1997 An efficient sol–gel reactor for antibody catalyzed transformations. *Chem. Mater* **9**, 2258–60

Shtelzer S, Rappoport S, Avnir D, Ottolenghi M, Braun S 1992 Properties of tryspin and of acid phosphatase immobilized in sol–gel glass matrices. *Biotech. Appl. Biochem.* **15**, 227–35

Slama-Schwok A, Avnir D, Ottolenghi M 1991 Photoinduced electron transfer mediated by a shuttling charge carrier between Ru and Ir trisbipyridyl complexes trapped in sol–gel glasses. *J. Am. Chem. Soc.* **113**, 3984–5

Tatsu Y, Yamashita K, Yamaguchi M, Yamamura S, Yamamoto H, Yoshikawa S 1992 Entrapment of glucose-oxidase in silica-gel by the sol–gel method and its application to glucose sensor. *Chem. Lett.* 1615–6

Tess M E, Cox J A 1999 Chemical and biochemical sensors based on advances in materials chemistry. *J. Pharmaceut. Biomed. Anal.* **19**, 55–68

Tsionsky M, Gun G, Glezer V, Lev O 1994 Sol–gel derived ceramic carbon composite electrodes: introduction and scope of applications. *Anal. Chem.* **66**, 1747–53

Tsionsky M, Lev O 1995a Electrochemical composite carbon ceramic gas sensors: introduction and oxygen sensing. *Anal. Chem.* **67**, 2409–14

Tsionsky M, Lev O 1995b Investigation of the mechanism and kinetics of co-porphyrin catalyzed oxygen reduction by hydrophobic carbon-ceramic electrodes. *J. Electrochem. Soc.* **142**, 2154–60

Turniansky A, Avnir D, Bronshtein A, Aharonson N, Altstein M 1996 Sol–gel entrapment of monoclonal anti-atrazine antibodies. *J. Sol–Gel Sci. Technol.* **7**, 135–43

Walcarius A J 1998 Analytical applications of silica-modified electrodes—a comprehensive review. *Electroanalysis* **10**, 1217–35

Wang J, Pamidi P V A 1997b Sol–gel-derived gold composite electrodes. *Anal. Chem.* **69**, 4490–4

Wang J, Pamidi P V A, Park D S 1996 Screen-printable sol–gel enzyme-containing carbon inks. *Anal. Chem.* **68**, 2705–8

Wang J, Pamidi P V A, Parrado C, Park D S, Pingarron J 1997a Sol–gel-derived cobalt phthalocyanine-dispersed carbon composite electrodes for electrocatalysis and amperometric flow detection. *Electroanalysis* **9**, 908–11

Wang J, Pamidi P V A, Rogers K R 1998 Sol–gel-derived thick-film amperometric immunosensors. *Anal. Chem.* **70**, 1171–5

Weetall H H 1996 Retention of bacteriorhodopsin activity in dried sol–gel glass. *Biosensors Bioelectron.* **11**, 327–33

Weigl B H, Wolfbeis O S 1994 Capillary optical sensors. *Anal. Chem.* **66**, 3323–7

Williams A K, Hupp J T 1998 Sol–gel encapsulated alcohol dehydrogenase as a versatile, environmentally stabilized sensor for alcohols and aldehydes. *J. Am. Chem. Soc.* **120**, 4366–71

Wolfbies O S, Reisfeld R, Oehme I 1996 *Sol–Gel and Chemical Sensors. Structure and Bonding*. Springer, Berlin, Vol. 85 pp. 51–98

Yang L, Saavedra S S, Armstrong N R, Hayes J 1994 Fabrication and characterization of low loss, sol–gel planar waveguide. *Anal. Chem.* **66**, 1254–63

Zheng L, Kulwinder F, Brennan J D 1998 Improving the performance of a sol–gel-entrapped metal-binding protein by maximizing protein thermal stability before entrapment. *Chem. Mater.* 3974–83

Zusman R, Rottman C, Ottolenghi M, Avnir D 1990 Doped sol–gel glasses as chemical sensors. *J. Non-Cryst. Solids* **122**, 107–9

D. Avnir, J. Blum and O. Lev
Hebrew University of Jerusalem, Israel

Reinforced Concrete

This article deals with reinforced structural concrete in its most general meaning. It addresses reinforced and prestressed concrete, as well as ferrocement, thin reinforced concrete composites, fiber-reinforced concrete, and hybrid combinations of these. A number of books and references are given in the Bibliography and should be consulted for any in-depth coverage. The Bibliography represents only a limited selection of US titles of the 1980s and 1990s and is not in any way complete or comprehensive. Its purpose is to provide the reader with a first link to additional information on any of the subjects discussed.

1. *Definition: Concrete Matrix and Reinforcement*

Concrete is a cement-based composite generally obtained by mixing a Portland cement binder, water, aggregates, and other additives. The aggregates may or may not be used; they generally comprise sand and gravel. Examples of additives include pozzolanic materials such as fly ash and silica fume, or simply very fine sand or clay. Additives may be inert or active. Depending on the aggregate content, different terms are used for concrete: a paste is made without aggregates; a mortar contains sand but no coarse aggregate (or gravel); a cement slurry is a very fluid paste containing a larger amount of water and/or some chemical additives to increase fluidity. For detailed coverage of concrete see *Portland Cements** and *Concrete as a Building Material**, and Mehta (1986) and Mindess and Young (1981). In this article the term concrete is used in the most general way to represent any one of the above-mentioned matrices.

Concrete is a durable construction material and has all the attributes suitable for constructed facilities.

From a mechanical viewpoint, all concrete matrices share a common property, i.e., they are weak in tension in comparison to their compressive strength and they have a low tensile strain prior to cracking. While formulas exists to estimate tensile strength based on compressive strength, as a first approximation the tensile strength of concrete ranges from about 7% to 10% of its compressive strength. Although very high strength concretes can be obtained, in most applications the compressive strength of concrete ranges from about 20 MPa to 60 MPa.

Since, in addition, concrete has a very low tensile strain capacity to failure it is brittle and it cracks. For all practical purposes in design, the tensile strength of concrete is ignored or assumed null.

To overcome the shortcoming of poor tensile resistance, reinforcement is added in concrete primarily in regions where tension is expected such as at the bottom of beams. However, in seismic regions where load reversals are expected, tension will occur at practically every portion of a section depending on the loading. So the reinforcement may be distributed along the section depth, as well as in two or three directions.

By its very definition and purpose, the reinforcement should have a tensile strength significantly higher than that of concrete. Also, when the reinforcement is not prestressed, a better performance or a more serviceable structural concrete is obtained when the modulus of elasticity of the reinforcement, E_r, is significantly larger than that of concrete, i.e., $E_r \gg E_c$.

When the reinforcement is embedded in concrete, both are generally assumed to be perfectly bonded to form a composite; this assumption has proved to be very effective in modeling the behavior of reinforced concrete members. In prestressed concrete the prestressed reinforcement may be bonded or intentionally unbonded; in the latter case, the reinforcement is anchored at the ends of the member.

2. Types of Reinforcement

The reinforcement in concrete structures can be either continuous or discontinuous or a combination of both (Fig. 1). The term "continuous" implies that the length of the reinforcement is of the same order as the length of the structural member, and generally larger than its least dimension. The term "discontinuous" generally implies a reinforcement significantly smaller than the length of the member and, except in thin reinforced sheets, also smaller that its least dimension; discontinuous reinforcement is described as fiber reinforcement.

2.1 Continuous Reinforcement

(*a*) *Steel reinforcing bars and prestressing tendons.* The most common form of reinforcement for concrete consists of round, surface-deformed steel reinforcing bars. However, welded wire fabric or grids can also be used. In the USA standard reinforcing bars are available in diameters ranging from about 9 mm to 57 mm and minimum specified yield strengths of either 276 MPa or 414 MPa. Wires are also available with minimum yield strengths of 483 MPa. Three types of steel tendons are used in prestressed concrete: wires, strands made with several wires, and bars. The most commonly used strands are made of seven wires (Fig. 2) and range in diameter from about 6.3 mm to 15.2 mm. They have a minimum specified tensile strength of 1860 MPa, and a yield strength of at least 85% of ultimate for normal relaxation and 90% of ultimate for low relaxation strand. Most common bars for prestressed concrete are deformed and range in diameter from about 19 mm to 35 mm. Their minimum specified tensile strengths are either 1000 MPa or 1104 MPa, and the minimum specified yield strength is 80% of tensile strength. Wires for prestressed concrete are generally less than 9 mm in diameter, and they can be used either individually, or as part of a cable assembly, or to make strands. Several strands can be assembled together to form cables for use in prestressing. Minimum properties of wires are similar to those of strands.

Reinforcing steel bars generally have an elastic plastic behavior in tension with a definite yield plateau (Fig. 3). Unlike reinforcing steels, prestressing steels do not show a definite yielding point; the onset of yielding is arbitrarily defined by standards. The yield strain of prestressing strands and wires is arbitrarily set by ASTM at 1%, while that of prestressing bars is set at 0.7%.

The modulus of elasticity of steel is almost independent of its strength. In lieu of information from the manufacturer, the following values can be used in design: 200 GPa for prestressing wires, 193 GPa for prestressing bars, and 186 GPa for prestressing strands.

Steel is the most common reinforcing material used with concrete. It has the advantages of being strong,

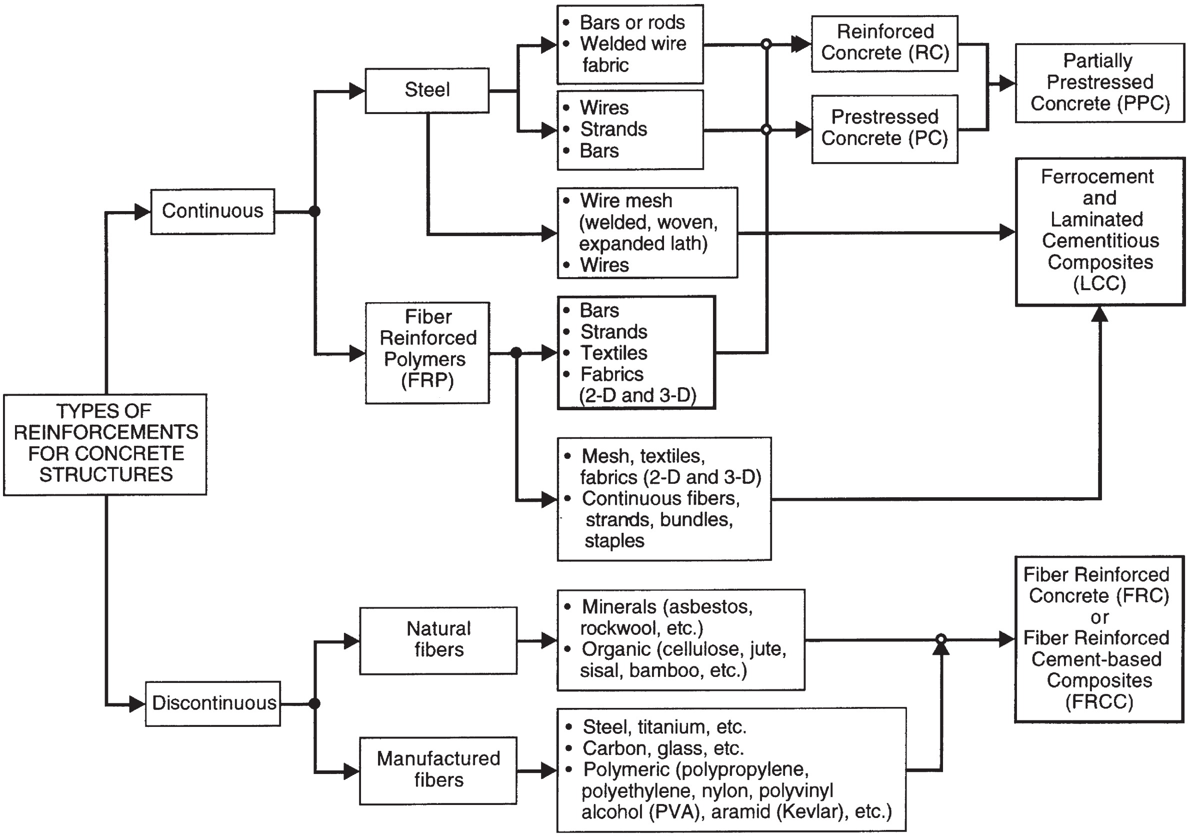

TYPES OF REINFORCEMENTS FOR CONCRETE STRUCTURES
Continuous
Steel
• Bars or rods
• Welded wire fabric
• Wires
• Strands
• Bars
• Wire mesh (welded, woven, expanded lath)
• Wires
Fiber Reinforced Polymers (FRP)
• Bars
• Strands
• Textiles
• Fabrics (2-D and 3-D)
• Mesh, textiles, fabrics (2-D and 3-D)
• Continuous fibers, strands, bundles, staples
Reinforced Concrete (RC)
Prestressed Concrete (PC)
Partially Prestressed Concrete (PPC)
Ferrocement and Laminated Cementitious Composites (LCC)
Discontinuous
Natural fibers
• Minerals (asbestos, rockwool, etc.)
• Organic (cellulose, jute, sisal, bamboo, etc.)
Manufactured fibers
• Steel, titanium, etc.
• Carbon, glass, etc.
• Polymeric (polypropylene, polyethylene, nylon, polyvinyl alcohol (PVA), aramid (Kevlar), etc.)
Fiber Reinforced Concrete (FRC) or Fiber Reinforced Cement-based Composites (FRCC)

Figure 1
Types of reinforcements used in concrete structures.

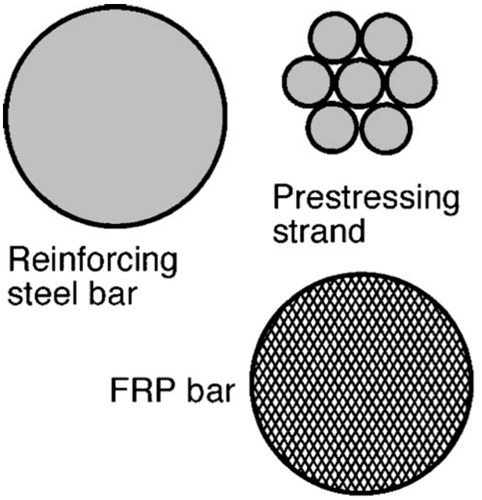

Reinforcing steel bar
Prestressing strand
Steel wire
FRP bar

Figure 2
Cross-section of typical reinforcing elements.

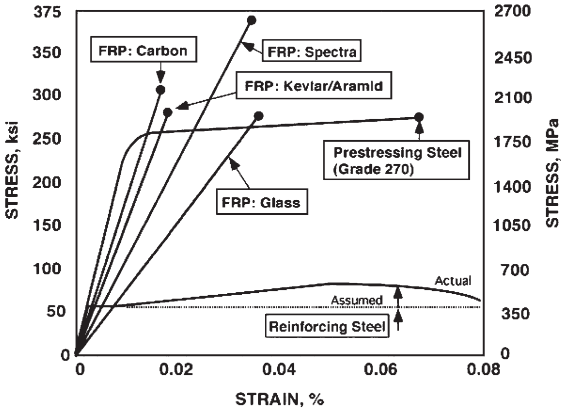


Figure 3
Typical stress–strain curves in tension of reinforcement materials.

ductile, and with a relatively high modulus of elasticity in comparison to concrete. However, it is prone to corrosion and is relatively heavy to handle in the field. These drawbacks can be overcome by using fiber-reinforced polymeric or plastic (FRP) reinforcements using high-performance fibers.

(*b*) *FRP reinforcements*. FRP reinforcements, made from carbon, glass, aramid, or other high-performance fibers embedded in polymeric matrices, in the form of bars, tendons, strands, and two- or three-dimensional meshes, are produced and tested by various institutes around the world especially in Japan, the USA, and Western Europe (ACI 1997). Examples of FRP reinforcements include tendons with trade names like Parafil, Arapree, CFCC, Polystal, Fibra, Leadline, and numerous others.

The use of FRP reinforcements is seen primarily as a means to avoid corrosion problems otherwise encountered in concrete structures when using conventional steel reinforcing bars or prestressing tendons. Moreover, their nonmagnetic properties make them ideal for special applications, such as radar stations and structures for magnetic levitation (mag-lev) trains. FRP reinforcements are suitable for use in new structures, as well as for repair or strengthening of existing structures, where their properties can be most advantageous. In the latter case, FRP sheets are adhesively bonded to the concrete surface to provide strengthening in bending, or shear, or both; they are also used to confine reinforced concrete columns to improve their seismic resistance.

Other advantages of FRP reinforcements, particularly in prestressing applications, include lower unit weight, ease in coiling and handling, good damping and fatigue behavior, and low relaxation losses. However, FRP materials are not without drawbacks which include high cost, low shear (transverse) strength, susceptibility to stress-rupture effects, and, most importantly, low ductility (i.e., no yielding). Indeed they show a linear elastic response in tension up to failure (that is considered a brittle failure) which leads to poor sectional and member ductility. Also, they have an elastic modulus lower than that of steel, and generally ranging from about twice (glass FRP) to about five times (carbon FRP) that of normal strength concrete. The reader is strongly advised to consult technical papers or texts on the subject prior to designing concrete structures with FRP reinforcements. The design suggestions made in this chapter are applicable to concrete structures with steel reinforcements, and cannot be applied to FRP reinforcements without special limitations.

2.2 Discontinuous Fiber Reinforcement

Short fibers used in concrete can be characterized in different ways. First, according to the fiber material: natural organic (such as cellulose, sisal, jute, bamboo, etc.); natural mineral (such as asbestos, rock-wool, etc.); and synthetic (such as steel, titanium, glass, carbon, polymers, synthetics, etc.). Second, according to their physical/chemical properties: density, surface roughness, chemical stability, nonreactivity with the cement matrix, fire resistance or flammability, etc. Third, according to their mechanical properties: tensile strength, elastic modulus, stiffness, ductility, elongation to failure, surface adhesion, etc.

Fibers are differentiated by a number of other important geometric characteristics related to their cross-sectional shape, length, diameter or equivalent diameter, and surface deformation. The cross-section of the fiber can be circular, rectangular, diamond, square, triangular, flat, polygonal, or any substantially polygonal shape. To develop a better bond between the fiber and the matrix the fiber can be modified along its length by roughening its surface or by inducing mechanical deformations. Thus fibers can be smooth, deformed, crimped, coiled, twisted, with end hooks, paddles, buttons, or other anchorages. Typical examples of steel fibers are shown in Fig. 4.

Most common steel fibers are round, have a diameter ranging from 0.4 mm to 0.8 mm, and a length ranging from 25 mm to 50 mm. Their aspect ratio, i.e., the ratio of length to diameter or equivalent diameter, is generally less than 100, with a common range from 40 to 80. The length and diameter of synthetic fibers vary greatly. Single-filament fibers can be as little as 10 μm in diameter such as for Kevlar, and as large as 0.8 mm such as with some polypropylene or poly(vinyl alcohol) (PVA) fibers. Generally, the aspect ratio of very fine fibers exceeds 100 while that of courser fibers is less than 100.

Due to the formulation of the mechanics of the composite, the fiber content in cement matrices is

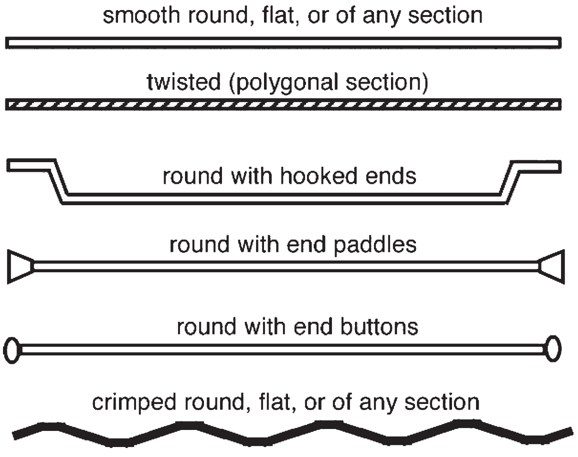

smooth round, flat, or of any section
twisted (polygonal section)
round with hooked ends
round with end paddles
round with end buttons
crimped round, flat, or of any section

Figure 4
Typical profiles of steel fibers used in concrete.

specified by volume fraction of the total composite. The fiber volume content of a typical fiber-reinforced concrete is less than about 2%. High-performance fiber-reinforced cement composites generally use fiber contents ranging from 2% to up to 15%, with a most common range between 2% and 6%. Because of fiber materials of different densities, the same volume fraction of fibers of different materials leads to different weight fractions of fibers. Fibers are purchased by weight. Typically a 1% volume fraction of steel fibers in normal-weight concrete amounts to about 80 kg m^{-3} of concrete, whereas a 1% volume fraction of polypropylene fibers amounts to about only 9.2 kg m^{-3}.

To be effective, fibers must have the following properties: (i) a tensile strength significantly higher than that of concrete (2–3 orders of magnitude higher); (ii) a bond strength with the concrete matrix preferably of the same order as or higher than the tensile strength of the matrix; and (iii) an elastic modulus in tension significantly higher than that of the concrete matrix, unless self-stressing is used through fiber reinforcement.

The mechanics of fiber reinforcement of concrete and other information on the applications of fiber-reinforced concrete can be found in Balaguru and Shah (1992), Bentur and Mindess (1990)), Naaman and Reinhardt (1996), Naaman (2000), and Reinhardt and Naaman (1992, 1999).

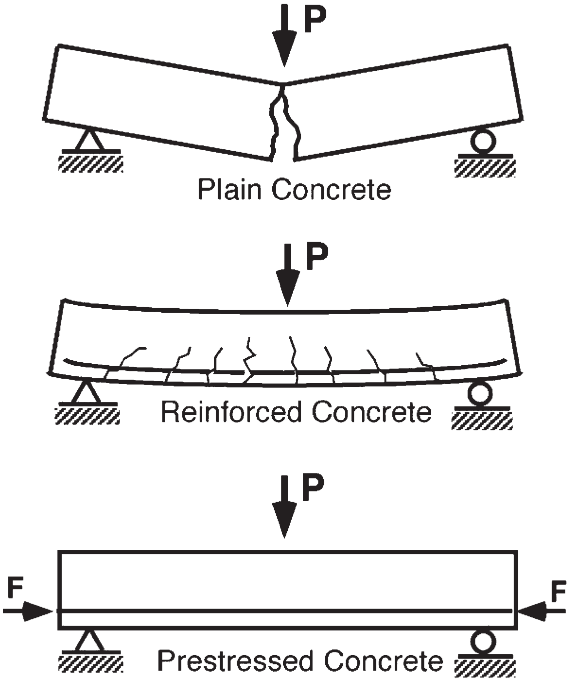


Figure 5
Response of plain, reinforced, and prestressed concrete to service loading.

3. Methods of Reinforcement of Concrete Structures

Generally all concrete structures are reinforced with some type of reinforcement. Nonreinforced concrete members are assumed to carry only small gravity loads or perform a noncritical, non-life-threatening load-carrying function. Examples include a slab on grade, a walkway, a low-height retaining wall, etc. All other applications belong to structural concrete where some type of reinforcement is used. It would be unsafe, for example, to use a nonreinforced concrete beam to sustain any external load including its own weight. Figure 5 illustrates the response of such a beam under service load and how reinforcement is placed to prevent its collapse. In reinforced concrete, bars are placed where tension and cracking are expected. In prestressed concrete, the tendons are stressed to apply a compressive force balancing tensile stresses and preventing cracking.

In concrete structures the reinforcement may be considered passive such as in reinforced concrete or ferrocement, or active such as in prestressed concrete. Active reinforcement implies that the reinforcement is stressed prior to application of external load to the structure.

Figure 6 illustrates the effects of prestressing for a concrete member subjected to tension. The reinforced concrete member will crack under load. When prestressing is used, the prestressed concrete member will not crack under load. This is because the uniform precompression stress induced by prestressing balances the tensile stress induced by external loads; the resulting stress remains a compressive stress throughout the member.

In reinforced concrete the bond is assumed to be perfect between the steel and the concrete. In prestressed concrete the prestressing tendons may be bonded or unbonded. In unbonded construction the prestressing force is transmitted to the concrete member at the end anchorages of the tendon only, as if it were an externally applied force. Unbonded prestressing tendons may be placed inside or outside the concrete section. In the latter case they are described as external tendons leading to external prestressing (Fig. 10). In its simplest definition, partially prestressed concrete is a combination of reinforced and prestressed concrete, i.e., a concrete member containing simultaneously reinforcing bars and stressed, prestressing tendons.

3.1 Various Types of Structural Concrete

Figure 7 schematically compares plain concrete with different forms of structural concrete beams

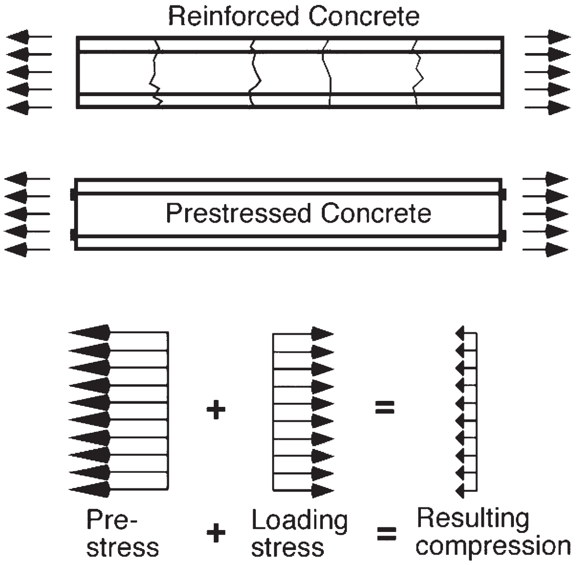

Reinforced Concrete
Prestressed Concrete
+
=
Pre-stress
+
Loading stress
=
Resulting compression

Figure 6
Internal stresses designed to balance external stresses in prestressing.

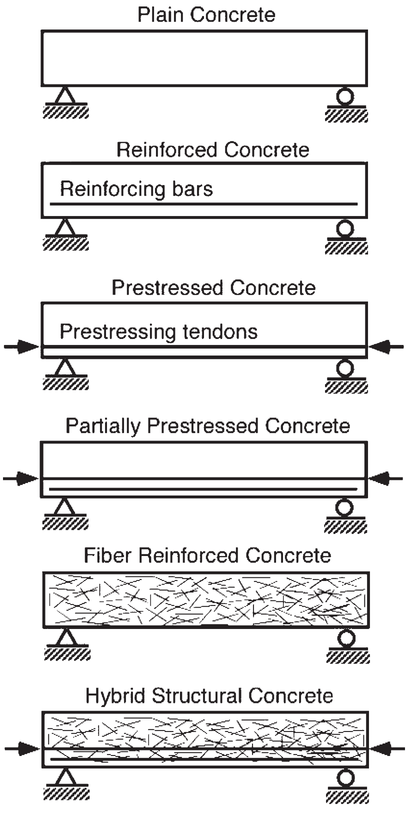

Plain Concrete
Reinforced Concrete
Reinforcing bars
Prestressed Concrete
Prestressing tendons
Partially Prestressed Concrete
Fiber Reinforced Concrete
Hybrid Structural Concrete

Figure 7
Various forms of structural concrete.

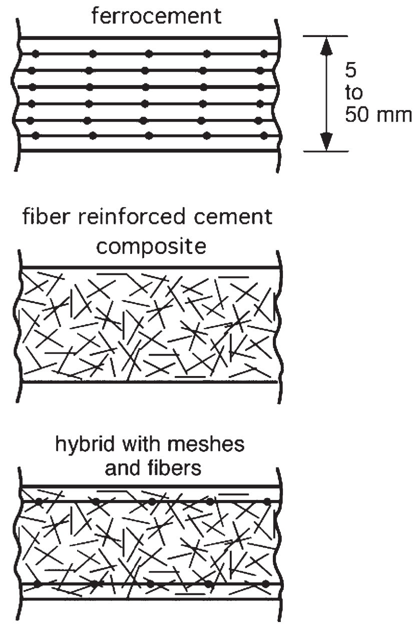

ferrocement
5
to
50 mm
fiber reinforced cement composite
hybrid with meshes and fibers

Figure 8
Typical sections of ferrocement and thin-sheet reinforced concrete products.

and provides the terminology used to designate each type of reinforcing system.

Reinforced and prestressed concrete structures can be built of practically any thickness. Nuclear power vessels can be more than 3 m thick, and dams can be as thick as a road is wide. However, it is not practical to build conventional concrete structures with a thickness of less than about 100 mm. Other forms of structural concrete such as ferrocement, fiber-reinforced concrete, or hybrid composites must be used (Balaguru and Shah 1992, Bentur and Mindess 1990, Naaman 2000).

Ferrocement is a type of thin-wall reinforced concrete commonly constructed of hydraulic cement mortar reinforced with closely spaced layers of continuous and relatively small size wire mesh (Fig. 8). The mesh may be made of metallic or other suitable materials. The fineness of the mortar matrix and its composition should be compatible with the mesh and armature systems it is meant to encapsulate. The matrix may contain discontinuous fibers or microfibers. The thickness of ferrocement is generally less than 50 mm, and in most applications less than 25 mm. The principles governing the design and analysis of ferrocement are similar to those used for reinforced concrete; however, a number of differentiating factors exists between them and should be evaluated prior to embarking on design and construction with ferrocement composites (Naaman 2000).

In the above paragraphs, it is assumed that continuous reinforcement is used. Short, strong, preferably stiff and ductile discontinuous reinforcing fibers can also be used with concrete matrices. The fibers can be added to the concrete matrix similarly to adding sand or other components; they are premixed with the concrete to lead to fiber-reinforced concrete. However, several forms of fiber-reinforced concrete exist where the fibers are not premixed. In some shotcreeting operations, the fiber (such as glass fiber) is fed separately to the nozzle where the matrix is projected under pressure. In slurry-infiltrated fiber concrete (SIFCON), the fibers are preplaced in the mold to form a three-dimensional network then infiltrated by a cement slurry. In slurry-infiltrated mat concrete (SIMCON), the fibers are first used to produce a three-dimensional porous mat which is then infiltrated by a cement slurry in a manner similar to SIFCON (Naaman and Reinhardt 1996, Reinhardt and Naaman 1992, 1999).

Fibers or microfibers can be added to the concrete matrix used in conventional reinforced concrete, prestressed concrete, or ferrocement. In this case a hybrid system or a hybrid composite is obtained. While fibers generally tend to improve a number of material and structural properties of concrete, the most predominant improvement is in the energy absorption capacity of the matrix or its fracture energy. Moreover, when used in the right proportion, their presence may improve the tensile and shear resistance of concrete, the bond between reinforcing bars and concrete, and the cyclic response (Balaguru and Shah 1992, Bentur and Mindess 1990).

4. Prestressing

4.1 Definition

Prestressing is the deliberate creation of permanent internal stresses in a structure or system in order to improve its performance. Such stresses are designed to counteract those induced by external loadings. Prestressed concrete is a mature structural material with a proven technology; a number of textbooks have been written about it (Collins and Mitchell 1991, Gerwick 1993, Libby 1984, Lin and Burns 1981, Naaman 1982, Nawy 1995a, 1995b, Nilson 1987).

The application of prestressing to concrete is in a way a natural result. Concrete is strong in compression and weak in tension. For design purposes its tensile resistance is discounted. Prestressing the concrete would produce compressive stresses, either uniform or nonuniform, which will counteract tensile stresses induced by external loading (Fig. 6). The original concept, initially developed by Eugene Freyssinet, attempted to counteract tensile stresses entirely, thus producing a crack-free material during service. However, it has since evolved to counteract only in part externally induced tensile stresses, thus allowing tension and controlled cracking in a way similar to reinforced concrete. This has led to what is called partial prestressing.

Partially prestressed concrete occupies the whole spectrum of the reinforcing range between fully reinforced and fully prestressed concrete. It can be considered a combination of both.

4.2 Prestressing Methods

Several methods and techniques of prestressing are available. However, except for chemical prestressing, most can be classified within two major groups: pretensioning and post-tensioning.

(*a*) *Pretensioning*. In pretensioning the prestressing tendons (wires, strands) are stretched to a predetermined tension and anchored to fixed bulkheads or molds (Fig. 9). The concrete is poured around the tendons, cured, and, upon hardening, the tendons are released. As the bond between the tendons and the concrete resists the shortening of the tendons, the concrete gets compressed. The prefix "pre" in pretensioning refers to the fact that the tendons are put in tension *prior* to hardening of the concrete.

Depending on the pretensioned structural elements produced, the profile of the tendons is either straight (Fig. 9), such as in hollow-cored slabs, or has one or two deflected points (also called draping or hold-down points) such as in bridge girders.

Pretensioning is the method mostly used for the production of precast prestressed concrete elements in the USA because it offers great potential for mechanization. Efficient long-line production techniques with casting bed lengths of up to 600 ft (182 m) where individual elements are cast end to end are preferred, as they require a single tensioning operation. Elements of standardized cross-sections are mass produced, yet customized by varying the length of each element and by placing inserts, holes, or blockouts for the mechanical or electrical distribution systems. Accelerated curing often permits early removal of the elements and daily reuse of the forms (24-hour production cycle). Excellent quality control and optimum use of labor and materials are achieved.

In order to stretch the tendons, hydraulic jacks are generally used. Another technique for stretching the

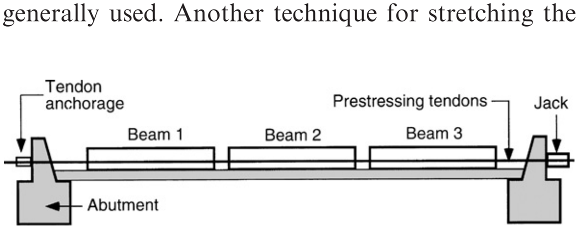


Figure 9
Typical pretensioning bed showing beams with straight tendons.

tendons is electrothermal prestressing. Here, high-strength deformed bars are heated by means of an electric current to between 250°C and 450°C, then placed and anchored at the ends of the molds or a pretensioning bed. Electrothermal prestressing is not used in the USA but has been used in the former Soviet Union and Eastern Europe. Its applications are limited.

(*b*) *Post-tensioning*. In post-tensioning the tendons are stressed and anchored at the ends of the concrete member *after* the member has been cast and has attained sufficient strength.

Commonly, a mortar-tight metal pipe or duct (also called sheath) is placed along the member before concrete casting. The tendons may have been pre-placed loosely inside the sheath prior to casting or could be placed after hardening of the concrete. After stressing and anchoring, the void between each tendon and its duct is filled with a mortar grout which subsequently hardens. Grouting ensures bonding of the tendon to the surrounding concrete, improves the resistance of the member to cracking, and reduces the risks of corrosion of the steel tendons.

This post-tensioning technique implies using what are commonly called "bonded tendons." If the duct is filled with grease instead of grout, bonding is prevented throughout the length of the tendon and the tendon force is transferred to the concrete member only at the anchorages. This leads to "unbonded tendons." Unbonded tendons are generally coated with grease or bituminous material and wrapped with waterproof paper or placed inside a flexible plastic hose; they are placed in the forms prior to concrete casting. When the concrete reaches sufficient strength, the tendons are stressed and anchored. They remain unbonded throughout their length and during service of the structure. Unbonded tendons can also be placed outside the concrete section leading to what is termed "external prestressing" (Fig. 10).

Although post-tensioning can be used in precast prestressed operations, it is most useful in cast-in-place construction where large building and bridge girders cannot be transported, and for customized structures which need tensioning on the job site. Its application in slabs (unbonded tendons) or unusual projects, such as nuclear power vessels, TV towers, long-span bridges, and offshore structures has become a must and will certainly continue to expand.

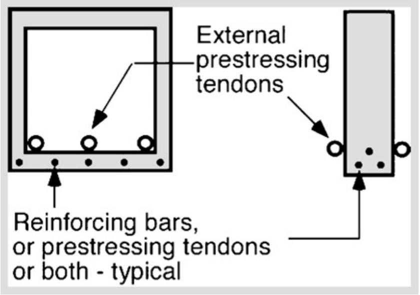

External prestressing tendons
Reinforcing bars, or prestressing tendons or both - typical

Figure 10
Illustration of external prestressing.

(*c*) *Chemical prestressing and self-stressing*. Pretensioning and post-tensioning represent two general methodologies to which most prestressing techniques belong. It was mentioned above that electrothermal prestressing is in effect a pretensioning method. Chemical prestressing, however, does not belong to any of the above. In chemical prestressing the tendons are placed untensioned in the forms and the concrete is poured. Due to the special expansive cement used, the concrete, instead of shrinking, expands after curing and during hardening. As the steel is bonded to the concrete, it stretches with it, thus undergoing tension and inducing compression in the concrete. Because stresses due to chemical prestressing are generally low, this technique has not been widely used.

There are "smart" materials that allow us to envision self-stressing with effective prestress levels that can be controlled much more accurately, and can be of much larger magnitude than is achieved with expansive cement matrices or electrical prestressing. Shape memory alloys (SMAs) and some special polymeric fibers possess the unique property of being able to be frozen temporarily in a particular state, then, with proper heat or radiation treatment, go back to their previous equilibrium state. Compared to electrical prestressing these materials do not need costly specialized electric equipment and do not create safety problems in the field. The treatment (heat or radiation) can be applied any time after hardening of the matrix instead of during its curing and hardening. The special reinforcement needed, such as SMA, can be factory produced, stored, shelved, placed in the composite, and triggered to recover its deformation (inducing prestressing) at any appropriate time.

Finally, it is conceivable to combine the previous two methods, i.e., to have a matrix that expands and a reinforcement that shrinks in order to produce the proper amount of self-stressing. The technique of self-stressing using shape memory materials was still at the basic research stage at the end of the 1990s (Naaman 2000, Reinhardt and Naaman 1999).

5. *Typical Behavior of Structural Concrete Beams Under Load*

When a reinforced concrete beam is loaded in bending it undergoes several stages of behavior. These

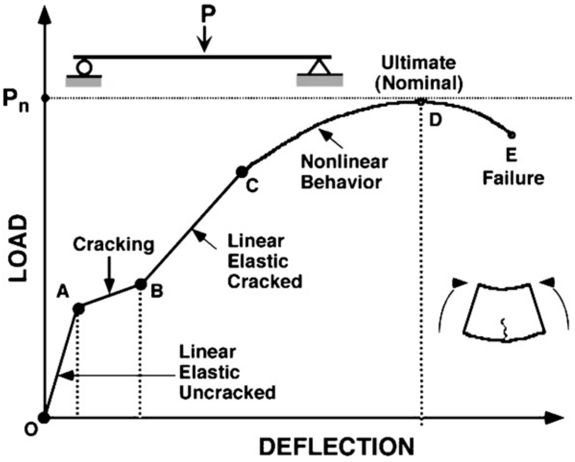


Figure 11
Typical moment vs. deflection response of a typical reinforced concrete beam.

help explain the basis for the different analysis and design approaches followed for reinforced concrete.

Consider the load (or moment) vs. deflection response of a typical beam (Fig. 11). Several portions of the curve can be identified.

(i) An initial portion (OA) where structural cracking does not occur.

(ii) A cracking region (AB) where cracks develop over a relatively short load range; their width increases thereafter with increasing load.

(iii) An almost linear portion (BC) where the concrete is cracked and both the concrete and the steel can be considered in their linear elastic range of behavior. The slope of this portion is smaller than that of the initial portion.

(iv) A region where the response of the concrete is nonlinear (CD) and/or where gradual yielding of the steel reinforcement occurs. This nonlinear region eventually leads to a maximum load (represented by point D) considered to be the nominal bending resistance.

(v) A post-peak portion (DE) at the end of which failure occurs either due to concrete failure in compression or due to failure of the extreme layer of reinforcement.

The analysis of stresses along each portion of the curve for the cases of reinforced concrete, prestressed concrete, and partially prestressed concrete is beyond the scope of this article since each commends a different procedure and a different set of equations. However, in predicting the nominal bending resistance (i.e., the maximum load for design purposes), a unified procedure can be followed for all reinforced and prestressed concrete members.

Indeed when a structural concrete section is subjected to bending in the cracked state, an internal couple develops with the compression force in the concrete and the tensile force in the steel reinforcement. The forces must remain equal. The compression force in the concrete balances the tensile force, whether it is resisted by reinforcing bars, prestressing tendons, or their combination. Since all steel reinforcements are assumed to yield, the initial prestress, if present, does not influence ultimate resistance. Figure 12 illustrates the section forces at nominal resistance. The corresponding detailed solution is given in Sect. 7.1.

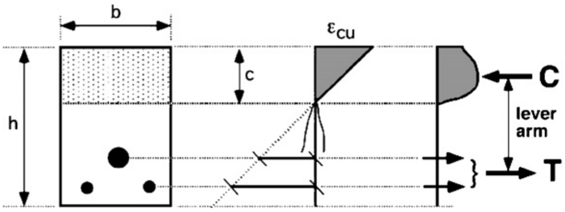


Figure 12
Typical model of a structural concrete section at nominal bending resistance.

6. Design Philosophy

The design philosophy generally adopted for concrete structures is based on satisfying a number of limit states where a limit state is defined as the boundary between acceptable and unacceptable performance. Limit states are generally classified in two categories, namely serviceability limit states and ultimate strength limit states. An allowable stress, a maximum crack width, or a limit value of deflection belong to the category of serviceability limit states. Collapse of the structure due to excessive bending or shear belongs to the category of ultimate strength limit states.

In the following formulation, the subscript u such as in M_u refers to ultimate and the subscript n such as in M_n refers to nominal. M_u is the required ultimate moment or the factored moment and M_n is the nominal moment resistance of the member, i.e., its maximum bending resistance. The factor ϕ is the capacity reduction factor or resistance factor and is defined below.

6.1 General Design Approaches

Design approaches applicable to all structural materials commonly use the following terminology:

USD: ultimate strength design
LSD: limit state design
LRFD: load and resistance factor design
WSD: working stress design
ASD: allowable stress design

WSD or ASD, also sometimes called the permissible stress design, are different terminologies for the same approach, in which maximum stress criteria are specified; i.e., an allowable stress is not to be exceeded. The USD, the LSD, and the LRFD imply essentially the same philosophy wherein collapse is considered under factored loads.

Note that WSD and USD may result in the same member design, but USD or LRFD generally allows for a more rational distribution of safety reserve.

6.2 USD, LSD, or LRFD

When the USD, LSD, or LRFD approach is applied, the general format is given by the formula:

$$U = \sum \gamma_i Q_i \leqslant \phi R_n$$

where i = type of load, i.e., dead load, live load, wind load, etc.; Q_i = nominal load effect; γ_i = load factor corresponding to Q_i; U = required or specified ultimate strength; $\sum\gamma_i Q_i$ = factored loads = required or specified ultimate strength U; R_n = nominal strength or nominal resistance; ϕ = resistance factor (or capacity reduction factor) corresponding to R_n; and ϕR_n = design strength or design nominal resistance.

The left-hand side of Eqn. (1) represents the required resistance which is computed from structural analysis based on assumed loads, and the right-hand side of Eqn. (1) represents a limiting structural capacity provided by the selected members. In USD or LRFD, the designer compares the effect of factored loads to the strength actually provided. The term design strength refers to the resistance or strength, ϕR_n, that must be provided by the selected member. The load factors, γ_i, and the resistance factor, ϕ, reflect the fact that loads, load effects (the computed forces and moments in the structural elements), and the resistances can be determined only to imperfect degrees of accuracy.

The required strength for design, $U = \sum\gamma_i Q_i$, is defined as the maximum (absolute value) of force obtained from the specified load combinations. The load factors and load combinations recognize that when several loads act in combination with the dead load, only one load takes on its maximum lifetime value at a time, while the other loads are at arbitrary point-in-time values. Examples of load combinations specified by the American Concrete Institute (ACI) Building Code Requirements for Reinforced Concrete (ACI 1995) are as follows:

$$\begin{aligned} &1.4D + 1.7L \\ &0.75(1.4D + 1.7L + 1.7W) \\ &0.75(1.4D + 1.7W) \\ &0.9D + 1.3W \end{aligned} \tag{2}$$

where D, L, and W represent either the loads themselves (dead load, live load, wind load) or the load effects (i.e., the forces or moments caused by the loads). The coefficients are the applicable load factors, γ_i.

Values of resistance (or capacity reduction) factor, ϕ, given by ACI 318 (ACI 1995) for reinforced concrete are:

$$\begin{aligned} &\phi = 0.9 \text{ for bending} \\ &\phi = 0.9 \text{ for axial load in tension} \\ &\phi = 0.85 \text{ for shear} \\ &\phi = 0.75 \text{ for axial load in compression, assuming spiral confinement} \\ &\phi = 0.70 \text{ for axial load in compression, assuming tie confinement} \end{aligned} \tag{3}$$

The factor ϕ accounts for uncertainties in materials properties, analysis approximation, etc.

When the USD approach is applied to bending, direct axial tension, or shear in reinforced concrete, the following formulation is obtained:

$$M_\mathrm{u} = \gamma_\mathrm{D} M_\mathrm{D} + \gamma_\mathrm{L} M_\mathrm{L} \leqslant \phi M_\mathrm{n} \tag{4}$$

$$N_\mathrm{u} = \gamma_\mathrm{D} N_\mathrm{D} + \gamma_\mathrm{L} N_\mathrm{L} \leqslant \phi N_\mathrm{n} \tag{5}$$

$$V_\mathrm{u} = \gamma_\mathrm{D} V_\mathrm{D} + \gamma_\mathrm{L} V_\mathrm{L} \leqslant \phi V_\mathrm{n} \tag{6}$$

where M is the bending moment, N the tensile axial load, V the shear force, and other notation is as above.

Note that the load and resistance factors are obtained from calibration using reliability analysis. Load factors are higher for loads with a higher degree of uncertainty. This is why the live load factor for reinforced concrete is higher than the dead load factor. However, resistance factors are lower for types of members or materials with a higher degree of uncertainty. This is why the capacity reduction factor, ϕ, for axial load is smaller than that for bending.

6.3 WSD or ASD

When the WSD or ASD approach is used, the following general formula applies:

$$\text{Actual stress} \leqslant \text{Allowable stress} \tag{7}$$

where absolute values of stresses are considered. The allowable stress can be obtained from:

$$\text{Allowable stress} = \frac{\text{Strength}}{\text{Safety factor}} \tag{8}$$

For bending, axial tensile load, or shear the actual stress induced by loading in an uncracked section is directly proportional to the applied load:

$$\text{Actual stress} = N/A_c \text{ for axial tensile load, } N \quad (9)$$

$$\text{Actual stress} = M/S \text{ for a bending moment, } M \quad (10)$$

$$\text{Actual stress} = VK/Ib \text{ for a shear force, } V \quad (11)$$

where S is the section modulus, K is the moment of area above the shear plane, and b is the width of the shear plane. Other notation is standard. Tensile and bending stresses in the cracked section in the linear elastic range of behavior can be similarly computed as described in Sects. 3 and 4.

Thus the allowable stress design criterion can be generally written as follows:

$$\sum \sigma_i \leqslant \left(\frac{1}{\text{SF}}\right)\sigma_\text{n} \quad (12)$$

where σ_i is the stress due to loading i, σ_n is the equivalent elastic stress at nominal resistance, and SF is a safety factor larger than one. The summation is done over the applied load combination of interest. Assuming, for instance, a combination of dead load and live load (often termed service load ($D+L$)) leads to:

$$\sigma_\text{D} + \sigma_\text{L} \leqslant \left(\frac{1}{\text{SF}}\right)\sigma_\text{n} \quad (13)$$

where σ_D is the stress due to dead load and σ_L is the stress due to live load.

Since, under elastic bending, stresses and moments are directly proportional, Eqn. (13) can be rewritten as:

$$M_\text{D} + M_\text{L} \leqslant \left(\frac{1}{\text{SF}}\right)M_\text{n} \quad (14)$$

where M_n is the nominal bending resistance. The safety factor is often taken equal to 2.5 in the WSD–ASD approach.

In comparing Eqns. (4)–(6) with Eqns. (13) and (14), it can be observed that in the USD–LRFD approach the effects of loads are multiplied by a factor (different for each load) and compared to the nominal resistance, while in the WSD–ASD approach the load effects are used without modification and the resistance is divided by a factor. The difference, however, lies in the fact that in the USD–LRFD approach each load has a different load factor, while in the WSD–ASD approach the same safety factor is applied to the effects of all loads. Figure 13 graphically illustrates the two approaches assuming a bending member.

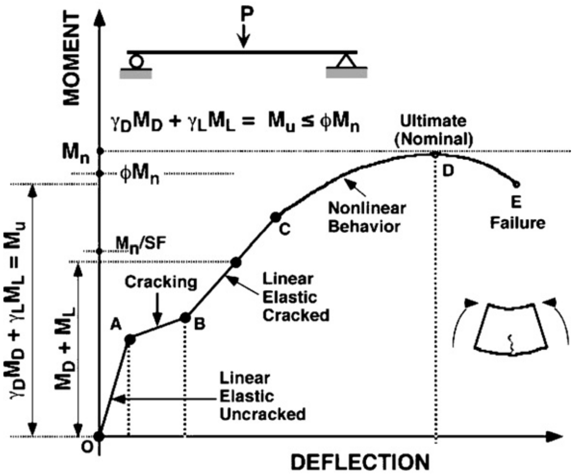


Figure 13
Illustrating the WSD and USD approach for a bending member.

7. *Analysis of Structural Concrete Sections at Ultimate: Nominal Bending Resistance M_n*

The purpose of the analysis at ultimate or maximum load is generally to determine the nominal moment resistance in bending (moment at ultimate behavior) of the section, assuming that cross-sectional dimensions, materials properties, and amounts of reinforcement are given. As pointed out earlier, the stress in the steel and in the concrete at ultimate are well outside their linear range of behavior. Although very accurate methods of analysis exists, a most efficient method used for everyday design is followed in the ACI Building Code (ACI 1995) where a rectangular stress block for the concrete is assumed and the reinforcing steel is considered to be at yield. Widely accepted assumptions for most codes are as follows:

(i) Plane sections remain plane under loading. Consequently it is assumed that a linear strain distribution exists along the concrete section up to ultimate load.

(ii) A perfect bond exists between steel and concrete. A direct result of this assumption is that any strain change in the steel due to an applied load is equal to the strain change in the concrete at the level of the steel due to the same load.

(iii) The tensile strength of concrete is neglected. Thus the point of zero stress represents the boundary between the cracked and the uncracked part of the section.

The ACI Building Code (ACI 1995) proposes the following additional assumptions (Fig. 14).

(iv) The limiting compressive strain of the concrete, ε_cu, is equal to 0.003, regardless of the strength of the concrete, the shape of the section, and the amount of reinforcement. It is also considered the same for normal weight and structural lightweight concrete.

(v) The maximum total force in the concrete compressive zone can be well approximated by considering

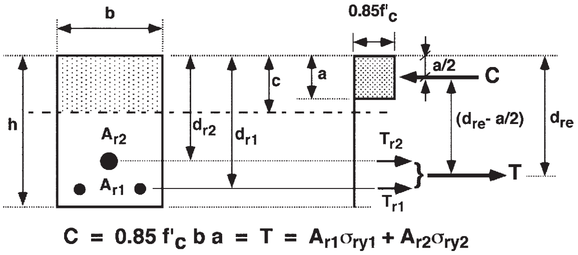


Figure 14
Internal resisting moment of the section at ultimate (or nominal bending resistance) assuming a concrete rectangular stress block in accordance with the ACI Building Code.

a uniform stress of magnitude $0.85f_c'$ over a rectangular block of width b and depth $a = \beta_1 c$, where c represents the depth of the neutral axis and β_1 is a factor defined by the ACI Code as follows (see Fig. 14):

$$\begin{aligned} &\beta_1 = 0.85 \text{ for } f_c' \leqslant 4000 \text{ psi (28 MPa)} \\ &\beta_1 = 0.65 \text{ for } f_c' \geqslant 8000 \text{ psi (56 MPa)} \\ &\beta_1 = 0.85 - 5 \times 10^{-5}(f_c' - 4000) \quad \text{for } 4000 \leqslant f_c' \leqslant 8000 \text{ psi} \\ &\beta_1 = 0.85 - 5 \times 10^{-2}(f_c' - 28)/7 \text{ for } 28 \leqslant f_c' \leqslant 56 \text{ MPa} \end{aligned} \quad (15)$$

While other shapes of compression stress block have been used in the technical literature, the rectangular stress block described above leads to the simplest prediction equations of bending resistance with little loss in accuracy.

7.1 Rectangular Section (or Rectangular Section Behavior)

Specialized texts treat in detail various cases of reinforcement such as the case of reinforced, prestressed, or partially prestressed concrete rectangular or flanged sections (T-sections) with and without reinforcement in the compression zone. Here only the case of a reinforced concrete rectangular section is illustrated and some simplifying assumptions are made to accommodate prestressed and partially prestressed concrete as well.

The following notation is used:

a = depth of rectangular stress block as per the ACI Building Code$a \times b$ = area of rectangular stress block for a rectangular section of width b

A_r = area of tensile reinforcement considered (if only one type)

A_{ri} = area of a particular type of reinforcement such as of reinforcing steel or prestressing steel, or both, or two types of reinforcing steel having different yield strengths

b = width of concrete section

c = distance from extreme compression fiber to neutral axis at ultimate

d_{r_i} = distance from extreme compression fiber to centroid of reinforcing material i

f_c' = specified design compressive strength of concrete or mortar

f_r = modulus of rupture or stress range in reinforcement

i = the type of reinforcement such as reinforcing steel, prestressing steel, or a reinforcing steel with different yield strength; in Fig. 14 the value of i is taken as 1 for the reinforcing steel and 2 for the prestressing steel; in many textbooks the subscript s is used for the reinforcing steel and p for the prestressing steel. The proposed notation allows for several layers of different steel reinforcements such as for ferrocement. It is assumed that all reinforcements are in the tensile zone

M_n = nominal bending resistance or ultimate moment

β_1 = factor defined in Eqn. (15)

ε_{mu} = ultimate strain of concrete matrix in compression (usually assumed to be 0.003 for mortar or concrete)

σ_{ryi} = specified yield strength of tensile reinforcement of type i

Referring to Figs. 12 and 14, the force equilibrium equation (of the internal couple) can be set by writing that the compressive force in the concrete compression block is equal to the total tensile force in the steel reinforcement, thus:

$$C = T \quad (16)$$

or

$$0.85 f_c' ba = \sum A_{ri}\sigma_{ryi} \quad i = 1, 2, 3\ldots \quad (17)$$

Thus, for instance, if we had a reinforced concrete section containing only one type of reinforcing steel in the tensile zone, Eqn. (17) would read:

$$0.85 f_c' ba = A_{r1}\sigma_{ry1} \quad (18)$$

For a section containing reinforcing steel and prestressed steel, Eqn. (17) would read:

$$0.85 f_c' ba = A_{r1}\sigma_{ry1} + A_{r2}\sigma_{ry2} \quad (19)$$

where the subscript 1 refers to reinforcing steel and 2 to prestressing steel.

Equation (17) allows us to compute the value of the depth of the stress block at nominal bending resistance:

$$a = \frac{\sum A_{ri}\sigma_{ryi}}{0.85 f_c' b} \quad i = 1, 2, 3\ldots \quad (20)$$

where the summation is taken for all different reinforcements, provided they have yielding behavior such as typical reinforcing steel or prestressing steel.

Generally in prestressed and partially prestressed concrete members with bonded tendons, the prestressed steel has a stress at nominal bending resistance somewhere between its yield strength and its ultimate strength. Since the yield strength of most prestressing steels exceeds 85% of their ultimate strength, assuming the stress to be at yield introduces a relatively small error. This is suggested by the author in Eqns. (17) and (19); however, various codes such as the ACI Building Code (ACI 1995) and the AASHTO bridge code (AASHTO 1994) recommend computing the stress in the prestressing steel using a prediction equation or a strain compatibility approach. The author believes that the approximation suggested above is sufficient in many cases of design and is generally on the safe side.

Once a is determined from Eqn. (20) the nominal bending resistance (i.e., the moment of the internal couple of Fig. 14) can be estimated from the equation of equilibrium of moments:

$$M_{\mathrm{n}} = \sum A_{\mathrm{r}i}\sigma_{\mathrm{ry}i}\left(d_{\mathrm{r}i} - \frac{a}{2}\right) \quad i = 1,\ 2,\ 3\ldots \tag{21}$$

which, for a section containing only reinforcing steel, leads to:

$$M_{\mathrm{n}} = A_{\mathrm{r1}}\sigma_{\mathrm{ry1}}\left(d_{\mathrm{r1}} - \frac{a}{2}\right) \tag{22}$$

For a section containing reinforcing steel and bonded prestressed steel Eqn. (21) gives:

$$M_{\mathrm{n}} = A_{\mathrm{r1}}\sigma_{\mathrm{ry1}}\left(d_{\mathrm{r1}} - \frac{a}{2}\right) + A_{\mathrm{r2}}\sigma_{\mathrm{ry2}}\left(d_{\mathrm{r2}} - \frac{a}{2}\right) \tag{23}$$

According to assumptions given in the ACI Building Code (ACI 1995), the depth of the equivalent concrete rectangular stress block, a, is related to the depth of the neutral axis at nominal bending resistance, c, by the following relation (Figs. 12 and 14):

$$a = \beta_1 c \tag{24}$$

or inversely:

$$c = \frac{a}{\beta_1} \tag{25}$$

in which the factor β_1 is defined above (Eqn. (15)) and a is from Eqn. (20).

In an analysis situation, the above equations of equilibrium (e.g., Eqns. (20) and (22) for reinforced concrete) have two unknowns, a and M_{n}, which can be solved. In a design situation, the value of A_{r} is unknown. By setting $M_{\mathrm{u}} = M_{\mathrm{n}}$ a third equation is introduced, thus allowing one to obtain a unique solution for a, M_{n}, and A_{r}. Hence, for design, A_{r} is obtained from Eqn. (22) where M_{n} is replaced by M_{u}/ϕ.

8. *Design for Bending by USD or LRFD*

8.1 *General Criteria for USD of Bending Members*

To satisfy strength design requirements in bending the following general conditions must be satisfied to insure strength and safety.

(i) Condition for strength:

$$M_{\mathrm{u}} \geqslant \phi M_{\mathrm{n}} \tag{26}$$

where M_{u} is the factored moment due to external loads, M_{n} is the nominal moment resistance, and ϕ is the strength reduction factor in bending (see above).

(ii) Condition for maximum reinforcement: this condition would guarantee some ductility for the member by forcing yielding of the steel prior to failure of the concrete.

(iii) Condition for minimum reinforcement: this condition insures that the amount of reinforcement is sufficient in order to warrant that failure/collapse does not occur just after cracking.

8.2 *Strength Criterion*

Whether an analysis, investigation, or design of a given section is carried out, once the nominal bending resistance, M_{n}, is calculated, the following design criterion must be checked (Eqn. (26)):

$$M_{\mathrm{n}} \geqslant \frac{M_{\mathrm{u}}}{\phi} \tag{27}$$

which implies for design that the nominal bending resistance $M_{\mathrm{n}} \geqslant M_{\mathrm{u}}/\phi$. When used at equality, this condition, combined with the two equations of equilibrium, allows for the computation of the required amount of reinforcement.

8.3 *Maximum Reinforcement*

Let us assume that a section satisfying Eqn. (27) has been determined. To satisfy criteria related to the maximum and minimum reinforcement, it is convenient to determine the depth, d_{re}, from the extreme compression fiber to the centroid of the tensile force in the reinforcement. Its value can be calculated from the following equation (Fig. 14) assuming all tensile reinforcement is at yield:

$$d_{\mathrm{re}} = \sum A_{\mathrm{r}i}\sigma_{\mathrm{ry}i}d_{\mathrm{r}i} \Big/ \sum A_{\mathrm{r}i}\sigma_{\mathrm{ry}i} \tag{28}$$

To guarantee ductile behavior or equivalently that excessive reinforcement is not used, the following general condition which is adopted in the AASHTO LRFD Bridge Design Specification (AASHTO 1994) is recommended:

$$\frac{c}{d_{\mathrm{re}}} \leqslant 0.42 \tag{29}$$

8.4 Minimum Reinforcement

To guarantee that failure does not occur at onset of first structural cracking, the ACI Building Code (ACI 1995) and the AASHTO LRFD Bridge Design Specification (AASHTO 1994) provide different requirements for reinforced and prestressed concrete structures. The author developed the following simple and unified condition which, while not given in any code, simulates very well the requirements given by ACI and AASHTO. It applies to all types of reinforced, prestressed, and partially prestressed concrete sections:

$$\frac{c}{d_{\mathrm{re}}} \geqslant \frac{0.033}{\beta_1} \tag{30}$$

in which β_1 is from Eqn. (15) and d_{re} is from Eqn. (28).

9. Reinforced Concrete Tensile Members

The analysis and design of reinforced concrete tensile members presents no mathematical difficulty. Since concrete is weak in tension and will crack under service loads, its tensile resistance is neglected in design; thus the entire tensile load is carried by the reinforcement. However, since concrete does have some tensile resistance, it is necessary to provide a minimum amount of reinforcement in tensile members in order to avoid their sudden brittle failure and collapse at onset of concrete cracking.

9.1 Analysis

Consider the tensile prism shown in Fig. 15. The following notation is adopted:

A_{c} = cross-sectional area of composite tensile member

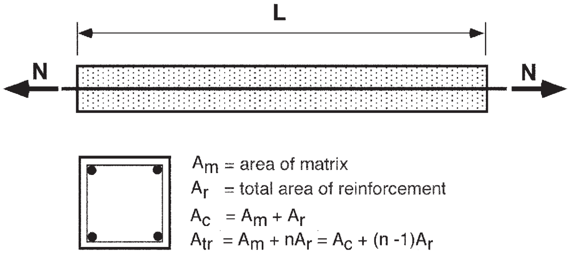

L
N
N
A_m = area of matrix
A_r = total area of reinforcement
A_c = $A_m + A_r$
A_{tr} = $A_m + nA_r = A_c + (n-1)A_r$

Figure 15
Assumed model of composite with load applied to the reinforcement.

A_{m} = cross-sectional area of matrix (concrete, mortar, paste)

A_{r} = cross-sectional area of reinforcement; here reinforcing steel only

A_{tr} = area of transformed section assuming all concrete

E_{m} = elastic modulus of matrix

E_{r} = elastic modulus of (longitudinal) reinforcement

$n = E_{\mathrm{r}}/E_{\mathrm{m}}$ = modular ratio

N = externally applied tensile load

N_{m} = tensile load resisted by the matrix

N_{r} = tensile load resisted by the reinforcement

ρ = reinforcement ratio; the definition for tensile members is different from that of bending members

u = subscript for ultimate

y = subscript for yield

ε_{c} = tensile strain in composite

ε_{m} = tensile strain in matrix

ε_{r} = tensile strain in reinforcement

σ_{c} = tensile stress in the composite

σ_{m} = tensile stress in the matrix

σ_{r} = tensile stress in the reinforcement

It follows from the above and from Fig. 15 that:

$$A_{\mathrm{c}} = A_{\mathrm{m}} + A_{\mathrm{r}} \tag{31}$$

$$\rho = \frac{A_{\mathrm{r}}}{A_{\mathrm{c}}} \tag{32}$$

$$n = \frac{E_{\mathrm{r}}}{E_{\mathrm{m}}} \tag{33}$$

For any applied external tensile load, N, the stress in the composite at any section is given by:

$$\sigma_{\mathrm{c}} = \frac{N}{A_{\mathrm{c}}} \tag{34}$$

The following assumptions are made.

(i) The load is applied to the reinforcement and then transferred to the concrete matrix; this is the same as if a through-matrix crack is assumed to exist at each end of the composite.

(ii) Under tensile load, cracking in the matrix occurs at an early stage when the stress in the matrix reaches its tensile strength, σ_{mu}.

The detailed analysis of stresses in the tensile composite in the precracking and postcracking stages and during the multiple cracking process can be found in other specialized texts (Naaman 1982, 2000). Here it will be sufficient to determine the necessary area of reinforcement needed to carry the tensile load with a sufficient level of safety reserve.

The stress in the composite at any section is given by Eqn. (34); stresses in either the matrix or the reinforcement, at a cracked section, are given by:

$$\sigma_m = 0 \tag{35}$$

and

$$\sigma_{\mathrm{r}} = N/A_{\mathrm{r}} \tag{36}$$

Thus

$$\sigma_{\mathrm{c}} = N/A_{\mathrm{c}} = \sigma_{\mathrm{r}}/\rho \tag{37}$$

or

$$\sigma_{\mathrm{r}} = \rho\sigma_{\mathrm{c}} \tag{38}$$

The minimum required value of reinforcement ratio for tensile members, ρ, is defined as that ratio of reinforcement for which collapse will not occur immediately after cracking. It can be shown that:

$$\rho_{\min} \geqslant \frac{1}{1 - n + \dfrac{\sigma_{\mathrm{ry}}}{\sigma_{\mathrm{mu}}}} \tag{39}$$

where σ_{mu} is the tensile strength of the matrix, σ_{ry} is the yield strength of the reinforcement, and n is the modulus ratio defined as $n = E_{\mathrm{r}}/E_{\mathrm{m}}$.

9.2 Design

For design the following equation can be used:

$$N_{\mathrm{u}} \leqslant \phi N_{\mathrm{n}} \tag{40}$$

where N_{u} is the factored tensile load and N_{n} is the nominal resistance in tension. The factored tensile load is defined above in Eqn. (5). The nominal tensile resistance in tension is taken as:

$$N_{\mathrm{n}} = A_{\mathrm{r}}\sigma_{\mathrm{ry}} \tag{41}$$

where A_{r} is the total cross-sectional area of tensile reinforcement and σ_{ry} is its yield strength.

In designing for the reinforcement, Eqn. (40) is taken at equality, leading to:

$$N_{\mathrm{n}} = \frac{N_{\mathrm{u}}}{\phi} \tag{42}$$

Thus the required area of reinforcement can be obtained:

$$A_{\mathrm{r}} = \frac{N_{\mathrm{n}}}{\sigma_{\mathrm{ry}}} = \frac{N_{\mathrm{u}}}{\phi\sigma_{\mathrm{ry}}} \tag{43}$$

It can be observed that for reinforced concrete tensile members the ultimate strength of the member is essentially the same as that of the reinforcement. This makes the computation of cross-sectional area of reinforcement very simple as shown above. However, one still needs to determine the area of the concrete tensile member or of the matrix. For this the following approach is suggested. Select a reasonable reinforcement ratio, ρ, such as from 2% to 4%. Thus:

$$A_{\mathrm{m}} \approx A_{\mathrm{c}} = \frac{A_{\mathrm{r}}}{\rho} \tag{44}$$

where A_{r} is from Eqn. (43). Make sure that $\rho > \rho_{\min}$ (Eqn. (39)).

It should be observed that while dimensioning the area of reinforcement and matrix is very simple, there is a need to detail the reinforcement in order to keep crack widths and stresses within acceptable limits. Crack width is proportional to the stress in the reinforcement. Also, for a given load, the higher the reinforcement ratio and the smaller the bar diameter, the smaller is the crack width. The design should also be checked for maximum displacement or elongation of the tensile member, which can be controlled by changing the area of the cross-section of concrete.

Once a preliminary cross-section is obtained, some iteration may be needed to fine-tune the results to select a safe, serviceable, and cost-effective member.

In prestressed concrete tensile elements, several criteria are used to determine the amount of reinforcement and the area of the concrete section. These include the ultimate tensile strength criterion similar to Eqn. (39) as well as criteria setting limits on compressive stresses, load to decompression, load to cracking, minimum reinforcement, and maximum deformation. Their treatment is covered in specialized texts (Naaman 1982).

10. Concluding Remarks

This article presents an overview of methods of reinforcement for structural concrete, which includes reinforced and prestressed concrete members as well as ferrocement and fiber-reinforced concrete. It describes the current limit state design philosophy for such members, and presents the underlying principles for the analysis and design of reinforced concrete beams and tensile elements. For additional details or for the design of other typical members such as columns or slabs, the reader is referred to more specialized publications on the subject listed in the Bibliography.

Bibliography

AASHTO 1994 *LRFD Bridge Design Specifications*. American Association of State Highway and Transportation Officials,Washington, DC

ACI Committee 318 1995 *Building Code Requirements for Reinforced Concrete and Commentary*, ACI 318-95 and ACI 318R-95. American Concrete Institute, Farmington Hills
ACI Committee 440 1997 *State-of-the-Art Report on Fiber Reinforced Plastic (FRP) Reinforcement for Concrete Structures*. American Concrete Institute, Farmington Hills
Balaguru P N, Shah S P 1992 *Fiber Reinforced Cement Composites*. McGraw-Hill, New York
Bentur A, Mindess S 1990 *Fiber Reinforced Cementitious Composites*. Elsevier, London
Collins M P, Mitchell D 1991 *Prestressed Concrete Structures*. Prentice Hall, Englewood Cliffs, NJ
Ferguson P H, Breen J E, Jirsa J O 1988 *Reinforced Concrete Fundamentals*, 5th edn. Wiley, New York
Gerwick B C Jr 1993 *Construction of Prestressed Concrete Structures*. Wiley Interscience, New York
Hassoun M N, Mamlouk M S 1998 *Design of Reinforced Concrete*. Addison-Wesley, Reading, MA
Leet K 1991 *Reinforced Concrete Design*, 2nd edn. McGraw-Hill, New York
Libby J R 1984 *Modern Prestressed Concrete*, 3rd edn. Van Nostrand Reinhold, New York
Lin T Y, Burns N H 1981 *Design of Prestressed Concrete Structures*, 3rd edn. Wiley, New York
MacGregor J G 1996 *Reinforced Concrete Mechanics and Design*, 3rd edn. Prentice-Hall, Englewood Cliffs, NJ
Mehta P K 1986 *Concrete: Structures, Properties and Materials*, 2nd edn. Prentice-Hall, Englewood Cliffs, NJ
Meyer C 1996 *Design of Concrete Structures*. Prentice-Hall, Englewood Cliffs, NJ
Mindess S, Young J F 1981 *Concrete*. Prentice Hall, Englewood Cliffs, NJ
Naaman A E 1982 *Prestressed Concrete Analysis and Design—Fundamentals*. McGraw-Hill, New York
Naaman A E 2000 *Ferrocement and Laminated Cementitious Composites*. Techno Press 3000, Ann Arbor, MI
Naaman A E, Reinhardt H W (eds.) 1996 *High Performance Fiber Reinforced Cement Composites*, HPFRCC 2, RILEM, No. 31 1996 E & FN Spon, London
Nawy E G 1995a *Prestressed Concrete: A Fundamental Approach*, 2nd edn. Prentice-Hall, Englewood Cliffs, NJ
Nawy E G 1995b *Reinforced Concrete: A Fundamental Approach*, 3rd edn. Prentice-Hall, Englewood Cliffs, NJ
Nilson A H 1987 *Design of Prestressed Concrete*, 2nd edn. Wiley, New York
Nilson A H, Darwin D 1997 *Design of Concrete Structures*, 12th edn. McGraw-Hill, New York
Paulay T, Priestley M J N 1992 *Seismic Design of Reinforced Concrete and Masonry Buildings*. Wiley, New York
Reinhardt H W, Naaman A E 1992 *High Performance Fiber Reinforced Cement Composites*, RILEM Publications 15. E & FN Spon, London
Reinhardt H W, Naaman A E 1999 *High Performance Fiber Reinforced Cement Composites—HPFRCC 3,* RILEM Proceedings, PRO 6. RILEM Publications, Cachan, France
Wang C K, Salmon C G, Wang C K, Anderson W 1998 *Reinforced Concrete Design*, 6th edn. International Textbook Co, New York

A. E. Naaman
University of Michigan, Ann Arbor, Michigan, USA

Roofing Materials

The basic function of roofing materials is to provide a pathway to direct storm water around the enclosed space. Functions such as shielding the structure from the effects of sun and wind are also important, but are secondary to the primary function of excluding and directing precipitation.

The entire roofing industry is divided into material and systems for low-sloped and steep-sloped roofing in recognition of the vital slope-to-drain function. Low-sloped roofs are usually defined as roofs having a slope of less than 25% (3 inches rise per foot (25 centimeters per meter)). Steep-sloped roofs are defined as roofs with a slope equal to or greater than 25%. The mean durability of selected steep and low-sloped roofs are shown in Figs. 1 and 2, drawn to the same scale. These data show the clear benefit of slope and how slope can greatly increase durability of the roofing by a factor of two or more.

1. *Steep-sloped Roofing*

Steep-sloped roofing became popular shortly after mankind left caves; it continues to be popular for residential roofing and is increasingly used on commercial and institutional buildings. In 1998 steep-sloped roofing accounted for 36% (9.5 billion dollars) of the total roofing market. Table 1 lists the 1998 market share (Kane 1999), and the mean service life and the life cycle cost from a 1995 roofing industry wide survey, for many types of steep roofing. The "other" category in Table 1 includes low-sloped roofing materials used at a slope of greater than 25%.

Approximately three quarters of the steep roofing material sold each year is for reroofing existing structures; the balance is used in roofing new construction. Of the reroofing portion, the old roofing materials are removed before the new materials are installed approximately 75% of the time; the old roofing materials are covered over in about 25% of the instances. All steep-sloped roofing materials are rain shedding rather than waterproof. As such, they must rely, to varying degrees, on the underlayment for additional water penetration resistance.

1.1 Asphalt–Glass Fiber Shingles

Asphalt–glass fiber shingles continue to dominate the steep-sloped shingle market in the USA due to their low life cycle cost and durability. The data in Table 1 is for "20 year" commodity shingles—a longer service life can be expected for the heavier and more costly shingles that prudence recommends. The total market volume in any year is directly proportional to population, while the local markets are dominated

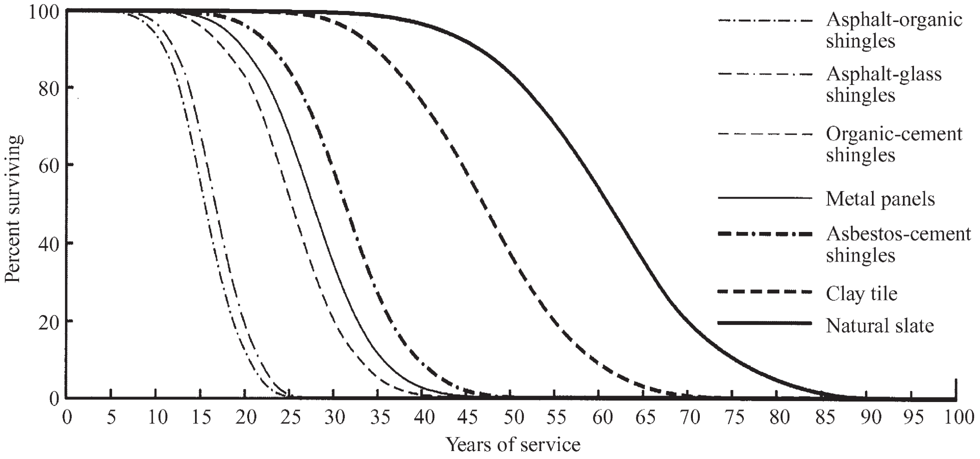


Figure 1
Mean durability of steep-sloped roofing systems.

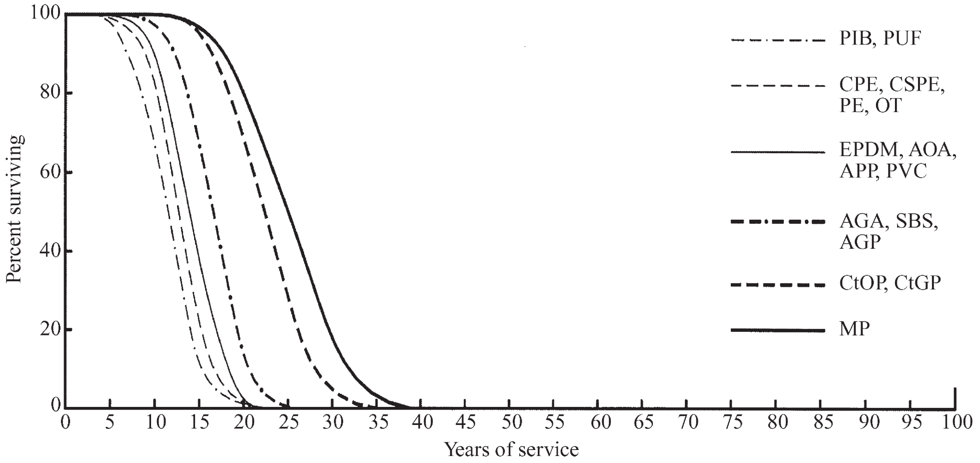


Figure 2
Mean durability of low-sloped roofing systems. AGA, asphalt coated glass fiber felts and asphalt built-up roofing; AGP, asphalt coated glass fiber felts and coal tar pitch built-up roofing; AOA, asphalt saturated organic felts and asphalt built-up roofing; APP, atactic polypropylene polymer modified roofing; CPE, chlorinated polyethylene; CSPE, chlorosulfanated polyethylene (Hypalon); CtGP, coal tar coated glass fiber felts and coal tar pitch built-up roofing; CtOP, coal tar saturated organic felts and coal tar pitch built-up roofing; EPDM, ethylene–propylene–diene terpolymer rubber; MP, metal panels; OT, other thermoplastic (thermoplastic polyolefin); PE, polyethylene; PIB, polyisobutylene; PUF, polyurethane foam; PVC, poly (vinyl chloride); SBS, styrene–butadiene–styrene block copolymer modified asphalt roofing.

by demographic changes and the state of the local economy.

These shingles are made by coating a non-woven glass fiber felt with hot asphalt that contains a substantial quantity (>50%) of fine mineral filler, embedding colored mineral granules in the top surface, dusting the back surface with a release agent such as talc or fine sand, often adding dots or strips of self-sealing adhesive, and die cutting the individual shingles. The shingles can weigh 3.5 to 6.2 kg m^{-2}, depending on the grade and design. The typical sales unit in the United States is a roofing square; that is, enough material to cover 100 ft^2 (9.3 m^2) of deck. The most popular current shingle design is the three-tab strip shingle in a single thickness, with overall dimensions of 12 × 36 in (305 × 914 mm). Laminated

Table 1
Steep-sloped roofing market share (1998), mean service life, and life cycle costs.

Roofing materials	Market share (%) New roofing	Reroofing	Mean service life (yrs)	Life cycle cost ($\$\text{ft}^{-2}\,\text{yr}^{-1}$)
Asphalt–glass shingles	47.1	48.8	18	0.20
Metal, architectural	17.3	8.4	28	0.74
Asphalt–organic shingles	11.2	14.0	18	0.20
Concrete tile	5.0	3.0	no data	no data
Clay tile	3.2	2.3	46	0.98
Metal, structural	2.7	1.8	24	0.36
Slate	2.7	2.0	60	1.09
Wood shingles	2.0	2.5	no data	no data
Cement shingles	0.3	0.4	no data	no data
Other	8.5	16.8	—	—

shingles are increasing in popularity because of the appearance of the added thickness they provide. Asphalt–glass fiber shingles are covered by ASTM standard D 3462.

Shingles are installed over an asphalt-impregnated felt underlayment with galvanized roofing nails. Field applied adhesive is used for very steep applications, and in areas that experience high winds. Specific application instructions are usually printed on each shingle wrapper.

Some manufacturers have introduced APP (atactic polypropylene) or SBS (styrene butadiene styrene block copolymer) polymer modified asphalt shingles. They have impressive flexibility and some have high wind resistance, but they have not yet been on the market long enough to demonstrate their performance.

1.2 Architectural Metal

Architectural metal such as copper, lead-coated copper, lead, and related alloys are used for roofing monumental buildings such as railroad stations, post offices, and capitol buildings. Copper continues to be the most popular, and can be gilded with gold leaf to provide the gold dome look of many capitols. Lead-coated copper or stainless steel is used where the roof runoff might stain the masonry or concrete walls.

The metal panels are installed over wood decking with nailed metal cleats, and locked joints may be soldered. The standing seam or flat seam panels are installed over an underlayment and use hidden fasteners.

1.3 Asphalt–Organic Felt Shingles

The asphalt shingles made on organic felt were precursor to today's asphalt–glass fiber shingles. They are manufactured by saturating a wood fiber felt with soft asphalt saturant before the sheet is coated, surfaced, and cut in a manner similar to that described for asphalt–glass fiber shingles. The higher asphalt content provided by the saturant reduces the fire resistance of the asphalt–organic felt shingles to an ASTM E108 Class C fire resistance (better than wood shakes but not as good as the Class A fire resistance of asphalt–glass fiber felt shingles). The asphalt–organic felt shingles are usually thicker, heavier, stronger, and more wind resistant (less brittle) than their asphalt–glass offsprings, and are preferred by many roofers in northern climates. It is not known if this preference is due to the conservatism of the roofers or the properties of the shingles, but the asphalt–organic shingles will be completely replaced as stronger and less brittle asphalt–glass shingles are introduced to the market.

1.4 Concrete Tile

Concrete tiles are a relatively recent ($\sim$1960) addition to the clay and concrete roof tile classification of roofing materials. They are primarily used in climates that experience few freeze–thaw cycles, where the high water absorption they demonstrate is of little consequence. Often overcoated with latex paints or whitewash to minimize dirt accumulation and fungus growth, concrete tiles are very popular in Florida. The tiles can be cleaned and recoated when they become dirty. Concrete tiles require a mechanically fastened, waterproof underlayment, and are attached to the underlayment with mortar, mechanical fasteners, or, more recently, plastic foam adhesive.

1.5 Structural Metal

Corrugated steel, aluminum, and similar metals are currently used for applications where they provide dual roles of waterproofing and structural support. With the advent of the floating roof concept, where factory-formed metal panels are allowed to expand and contract freely within their restraints, structural metal roofing became practical. Structural metal roofing is lightweight, can include any desired quantity of insulation, and is economical to install. It presents problems with odd shaped buildings, roofs that require many penetrations, and many of the details at penetrations and perimeters can be very challenging. On steep slopes structural metal roofing is water-shedding. Low-sloped structural metal roofs rely on factory-applied sealants or edge gaskets to be watertight; these seals may fail at ice dams or other severe exposures to ponded water.

1.6 *Natural Slate*

Natural slate shingles have the longest service life. As with any natural product, the local quality can vary significantly. Laboratory tests such as in ASTM Standards C217 and C406 should be applied to each source of slate. Individual slates are usually checked by an experienced slater as they are installed. It is important to use stainless steel or other fasteners with a long service life because the service life of the assembly might otherwise be limited by the life of the fasteners.

1.7 *Wood Shingles*

The wood shingle appearance is admired by many because of its rustic, rough texture; this appearance has been copied with many other materials in varying degrees of success. Its durability is dependent on the wood species used and the exposure. Wood shingles have a relatively low fire resistance and cannot be recommended where severe fire exposure may be present.

1.8 *Cement Shingles*

Many of the original cementitious shingles were reinforced with asbestos fibers and have a history of excellent performance. Since asbestos-containing products have fallen out of environmental favor, wood-fiber-reinforced cement shakes and shingles have been introduced. Without the long history of the asbesto-reinforced products, some of these wood-reinforced products are performing well, while others have failed to live up to the claims made by their manufacturers.

2. *Low-sloped Roofing*

Low-sloped roofing represented approximately 65% (~\$17 billion) of the 1999 roofing industry sales. About 25% of the sales represent roofing new construction; 75% was used for reroofing. In reroofing, the old roof was torn off ~65% of the time. The old roof was covered over with new material in ~35% of the reroofing.

Table 2 lists the market share, mean service life, and life cycle costs associated with many of the low-sloped roofing systems currently available. This is not a complete list because new types of systems are continuously being introduced. The "other" category in Table 2 includes many of these new materials and materials intended for steep-sloped applications.

2.1 *EPDM (Ethylene–Propylene–Diene Rubber)*

EPDM roofs are widely used in the northeastern part of the USA. They are used less frequently in the mid- and far-western states. EPDM is usually supplied in large sheets, making it ideal for large roof areas with few penetrations. EPDM is available in several thicknesses, in both reinforced and unreinforced, and in FR (fire-retardant) and regular grades. The rubber

Table 2
Low-sloped roofing market share (1998), mean service life, and life cycle costs.

Roofing materials	Market share (%) New roofing	Market share (%) Reroofing	Mean service life (yrs)	Life cycle cost ($\$$ft^{-2} yr^{-1})
EPDM	25.8	18.4	14	0.33
Asphalt built-up	17.5	1.4	17	0.31
SBS modified	10.6	11.1	16	0.34
APP modified	7.8	12.4	14	0.34
Metal, architectural	6.8	2.6	25	0.36
PVC	4.8	4.2	14	0.36
Sprayed polyurethane	2.4	2.3	12	0.42
TPO	2.3	1.8	13	0.37
Coal-tar built-up	2.3	2.8	22	0.30
CSPE	1.3	1.1	13	0.38
Other	18.4	23.9	—	—

sheets can be glued to the substrate with contact cements, mechanically fastened, or ballasted with an overburden of stones or concrete pavers. Currently, uncured butyl rubber tapes are used to seal the joints between the sheets. Reinforced EPDM is usually used for mechanically attached systems. EPDM has a moderate service life. It is attacked by oils and greases.

2.2 Asphalt Built-up Roofing

Asphalt membranes using glass fiber felts are the most popular of the built-up roofs; they have largely replaced their cousins made with organic felts. Properly applied, asphalt–glass built-up roofs are in the class of low-sloped roofs, with the longest service life. Three to four felt plies alternate with layers of asphalt; hence the built-up roof name. Asphalt built-up roofs should not be used in exposure to organic acids, and the odor of hot asphalt can be objectionable to sensitive individuals.

2.3 SBS (Styrene–Butadiene–Styrene Block Copolymer) Polymer-modified Asphalt

SBS polymer-modified asphalt membranes usually have two plies of coated felt. The top ply must be granule surfaced. The sheets are very flexible and are as extensible as their reinforcing allows. The sheets are usually installed in hot asphalt or cold process adhesives. SBS cap sheets are sometimes used over an asphalt–glass built-up membrane (composite construction) for difficult exposures. Excellent long-term exposures are reported from northern European countries.

2.4 APP (AtacticPolypropylene) Polymer-modified Asphalt

APP polymer-modified membranes use blends of polypropylene to modify the coating asphalt on sheets that are usually applied with a torch to melt the cap sheet onto a base sheet. Protective coatings are recommended to maximize the membrane's service life. Excellent weathering performance is reported from southern Europe.

2.5 Architectural Metal

Copper, lead-coated copper, stainless steel, and similar metals are used on low-slope applications. These low-sloped applications differ from the steep-sloped applications in that all the seams must be soldered waterproof, not merely bent to be water shedding. Systems are offered with gaskets and sealant-filled joints, but these should only be used on steep slopes because it is unwise to rely on sealant joints for water tightness.

2.6 PVC (Poly(Vinyl Chloride))

Reinforced PVC sheets are resistant to most chemicals; the seams between sheets are heat sealable, to provide a membrane where the integrity of each seam can be checked. The narrower (one meter) wide sheets are ideal for roofs with many penetrations. The PVC sheets are installed with mechanical fasteners, fully adhered, or ballasted. Some manufacturers offer waterproofing sheet systems. The PVC sheets are very slippery when wet.

2.7 Sprayed Polyurethane

Polyurethane foam is sprayed over the surface of the prepared substrate. The polyurethane foam is coated with acrylic, urethane, or silicone rubber coatings. Roofing granules may be broadcast into the wet coating to improve weatherability. Uncoated foam "rusts" (forms a brown dusty layer). The quality of the installed system is very dependent on the skill of the applicator, and "bird baths" (ponded water) should be avoided. This system is useful for domes and other roofs with irregular shapes; it has the highest life cycle cost of any of the other systems listed in Table 2.

2.8 TPO (Thermoplastic Polyolefin)

TPOs are a recent addition to the US market; they are widely used in Europe—principally in an inverted roof or a waterproofing application where they are protected from sunlight. The correct balance between fire-retardant and ultraviolet resistance additives is said to be critical. The principal olefins used are blends of polypropylene and polyethylene. These membranes are of great interest because it may be possible to recycle them into new membranes after they have outlived their service life.

2.9 Coal-tar Pitch Built-up Roofing

Coal-tar pitch membranes using coal-tar saturated organic felts is one of the oldest systems listed in Table 2. These systems perform very well historically, but are limited to very low slope applications and must be gravel surfaced because of the low viscosity of coal-tar pitch at normal rooftop temperatures. Coal-tar systems are prime candidates exposed to organic acids, such as the roofs on milk processing plants, or soup-canning factories. These membranes frequently contain four felt plies. Some people find the mothball odor evident during the roofing process objectionable; it is a carcinogen.

Coal-tar pitch membranes using compatible bitumen glass fiber felts are relatively new to the field. Some of these membranes made with asphalt-coated glass felts and coal-tar pitch have performed poorly because of incompatibility between the asphalt and the pitch. Additional problems have been reported when the glass fiber felts are too porous for the low viscosity pitch, permitting the felts to float upward before the gravel surface is applied, or allowing the felts to sink to the bottom of the membrane under the weight of the gravel surfacing. Standards have not been set for felt porosity.

2.10 CSPE (Chlorosulfanated Polyethylene)

CSPE membranes are often a blend of CSPE (Hypalon®) with PVC, or other thermoplastic polymers to permit the sheets to be heat sealable. The CSPE membranes have been decreasing in popularity because of performance difficulties.

3. *Summary*

The number of roofing materials on the market has been constantly increasing, and shows little sign of decreasing in the future, although many of the current systems may fall out of favor and disappear when competitive pressures or relatively poor life cycle performance become evident.

Bibliography

Cash C G 1997 The relative durability of low-sloped roofing. In: *4th Int. Symp. Roofing Technology*. National Institute of Standards and Technology and National Roofing Contractors Association, Gaithersburg, MD, pp. 119–24

Cash C G 2000 Estimating the durability of roofing systems. In: Evans J D, Ketola W D (eds.) *Durability 2000: Accelerated and Outdoor Weathering Testing. ASTM STP 1385 American Society for Testing and Materials*. West Conshohocken, PA

Kane K 1999 Another strong year for the roofing industry. *Professional Roofer* **29** (4), 14–7

C. G. Cash
Simpson Gumpertz and Heger, Inc., Arlington, Massachusetts, USA

Rubber Tires

The pneumatic tire is the most complex composite product in mass production today. It consists of a variety of materials such as natural rubber (NR), synthetic elastomers such as polybutadiene (BR), textiles such as nylon and polyester, and a range of different brass-coated steel wires. Since the 1960s, tires have undergone a major evolution. In the 1960s they were typically of a bias construction and required use of an inner tube for air retention. Bias or diagonal ply tires describe a tire construction where the ply cords extend to the beads, and where the ply cords are laid at an angle substantially less than 90° to the centerline of the tread or the direction of the beads cables (Fig. 1). Typically in this time period automobile tires reached up to 20 000 miles, after which they were replaced. In the 1970s, the radial tire emerged with significant improvements in mileage, fuel economy, and safety. Radial tires have a structure where the casing ply cords extend to the beads and are laid at a 90° angle to the centerline of the tread. The casing is stabilized by an essentially ridged belt system. This was followed by the introduction of the tubeless tire, which used an airtight membrane inside the tire, and formed an airtight seal between the tire and the rim through compressed fitting of the tire bead on an inclined steel wheel. "Ultralow-aspect ratio" tires are now emerging where the sidewalls of the tires are considerably shorter, allowing improvements in vehicle stability, rolling resistance, and, in the case of commercial trucks, cargo volume. They are also providing a basis for the development of "run flat" technology, where tires can operate at near zero inflation pressure.

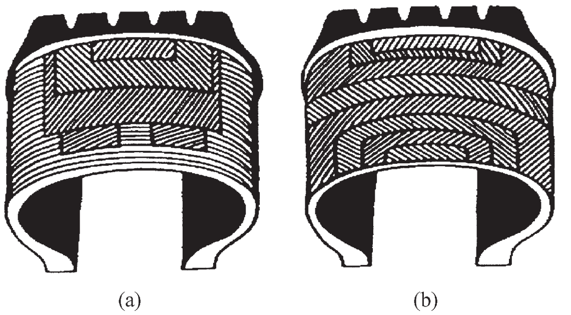

(a) (b)

Figure 1
Tire construction, (a) radial and (b) bias types.

Tires fall into essentially eight categories, based on the mission profile for the product. These are: racing and sports vehicles; passenger or consumer vehicles; light truck and sports utility vehicles (SUV), where gross vehicle weights do not exceed 7250 kg; commercial trucks; farm applications; earthmoving equipment, where gross vehicle weights can approach 300 tons; aircraft; and finally nonpneumatic specialty tires for applications such as fork lift trucks. All tires, whether they are designed for something as simple as a bicycle or as complex as a large commercial aircraft, must meet a fundamental set of performance functions:

(i) provide load carrying capacity;

(ii) provide cushioning, damping, and minimum noise and vibration;

(iii) transmit driving cornering, steering, and braking torque;

(iv) resist abrasion;

(v) have low rolling resistance; and
(vi) be durable and safe through the expected service life of the product.

1. Basic Tire Design

Figure 2 illustrates the components of a typical commercial truck tire used on vehicles which would have a gross vehicle weight (GVW) up to 40 tons. In summary, the components can be described as follows.

1.1 Tread

Tread is the component of the tire that is in contact with the road. It must show wear resistance, good traction characteristics, fuel economy, and resist service-related damage. A base underneath the tread is designed to dissipate heat and ensure good adhesion between the tread and the tire casing. Automobile tires tend to contain blends of styrene butadiene rubber (SBR) and polybutadiene (BR) rubber, oils, carbon black (up to 45% by weight), and silica. Heavy-duty truck tire treads tend to contain high levels of natural rubber (NR) and approximately 30% by weight of fine particle size carbon black. Other tires such as for aircraft, farm equipment, or earthmoving equipment can contain blends of NR, BR, and SBR.

1.2 Sidewall

The sidewall protects the tire from impact and curb scuffing. It also provides long-term weathering protection and casing durability. It is typically a blend of NR and BR and contains carbon black, high levels of antioxidants, and antiozonants.

1.3 Bead Region

The bead region consists of a number of components, such as the bead cable which locks the tire onto the wheel or rim, apex or bead filler, chipper wire which protects the ply from rim damage, and the chafer, which is the abrasion-resistant rubber component in contact with the rim.

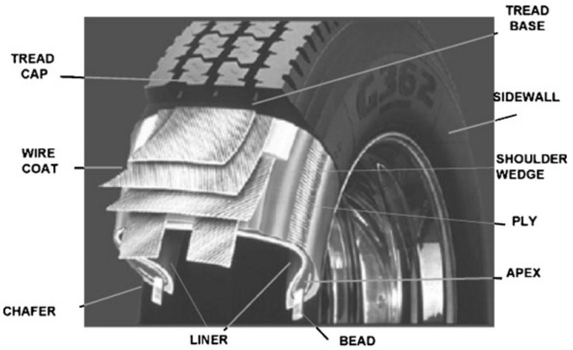


Figure 2
Radial truck tire components.

1.4 Plies

Consisting of steel or fabric, plies extend from bead to bead, and serve as the primary reinforcing member on the tire. They provide strength to the tire to enable high-pressure air retention. Fabric plies can be made from either poly(ethylene terephthalate) (PET), nylon-6, or nylon-66.

1.5 Belts

Belts consist of layers of steel wire and fabric forming a hoop under the tread. They restrict deformation of the casing and provide rigidity to the tread region, thereby allowing improved wear and handling performance, better damage resistance and protection of the ply cords. Belts are composed of either steel wire, PET, nylon-66, or nylon-6 coated with a NR based compound.

1.6 Shoulder

This is a wedge, which is a stiff component composed of compounded natural rubber, designed to provide stability of the belts and tread.

1.7 Liner

The liner is a membrane consisting of a compounded bromo-isobutene-isoprene rubber (BIIR), or other derivatives of this polymer or blends of butyl (IIR), halobutyl, and NR polymers, and whose function is to retain compressed air inside the tire.

2. Tire Engineering

Tire section width is the section width of the tire from sidewall to sidewall, illustrated in Fig. 2. The section height is the distance from the bottom of the bead of the tire to the top of the tread. These two parameters are important in describing the tire size seen on the sidewall. For example, the tire size designation LT235/75R15 denotes:

LT	light truck
235	approximate section width in millimeters
75	aspect ratio or the ratio of section width to section height
R	radial construction
15	nominal rim diameter in inches

Other sidewall markings include construction and manufacturing information, speed rating, and load carrying capability. Reference should be made to the

Tire and Rim Association Inc. Year Book, or the European Tyre and Rim Technical Organization (ETRTO) *Standards Manual*, for a full description of tire nomenclature and rim specifications for a given tire size.

The tread pattern influences the ability of the tire to transmit steering, braking and driving forces while operating in a broad range of highway and off-road surfaces. Reference should be made to Clark (1982) for a description of tread pattern design and influence on wear, traction, vehicle handling, tire damage resistance, and tire rolling resistance.

3. Tire Reinforcements

A tire is a textile/steel cord/rubber composite, where the steel and fabric cords reinforce the rubber compounds and are the primary load-bearing structure within the tire.

Polyester (PET) is used as the ply cord in automobile and light truck tires. It is also used for belts of fabric reinforced tires such as farm tires. The selection and strength of a polyester cord are defined by the tire performance requirements. These in turn are determined by the mission profile of the vehicle for which the tire has been designed.

Nylon-6 (polycaprolactam) and *nylon-66 (poly(hexamethylenediamine) adipamide)* are used in the plies of commercial tires (farm, aircraft, bias truck) and also as an overlay in high-performance passenger and light truck tires. The overlay sits over the steel belts and serves to hold the belts in place while the tire is operating at high speeds. Nylon-66 is preferred over nylon-6 due to its higher melting point (265 °C vs. 225 °C). Nylon has good fatigue resistance, durability, and tenacity. However, it tends to have poor dimension stability.

Steel cords used in tires are brass coated. Steel rod is first drawn down to a diameter of 1.2 mm and the brass plating applied. The filaments are then drawn down to their final diameter, which could range from 0.15 mm to 0.45 mm depending on the application for which the cord is needed. Steel cord construction terminology such as 3 + 6x.22 describes a cord of three filaments, wrapped with six additional filaments with a diameter of 0.22 mm. Steel cords used for plies have been designed for fatigue resistance. Belt wires have greater rigidity for stiffness. Ford and Charles (1988) have a comprehensive discussion on steel cord design and application.

Cotton and rayon are no longer used in tires to any significant extent. Aramid is used in special applications such as in aircraft tires. New technology cords under development include hybrid cords of nylon, aramid, and polyester, very high tensile strength steel cords, and new polyesters such as PEN (poly (ethylene-2,6-naphthalene dicarboxylate)), POK (polyethylene ketone), and carbon fibers.

Cord construction such as 1400/1/3 defines the linear density of base yarn, in decitex (1400), the number of yarns in the ply (1), and number of plies in the cord (3). The direction of twist on the filaments or yarns in a cord can be described as in the "S" direction, which is clockwise around the central axes of the tire cord, or "Z" twist, which is counterclockwise around the central axes (Bhakuni *et al.* 1997).

4. Tire Compounds

Compounding ingredients fall into one of five general categories: (i) polymers, which are typically natural rubber (NR), styrene butadiene rubber (SBR) and polybutadiene (BR); (ii) fillers such as carbon black, silica, clays, and calcium carbonate, though the latter two are used more in rubber products for industrial applications; (iii) protectant systems consisting of antioxidants, antiozonants, and waxes; (iv) a vulcanization system; and (v) special purpose chemicals such as processing aids, oils, and resins. Product performance objectives define the initial selection of materials. These materials must be benign with respect to environmental concerns, must facilitate processing in production plants, and be cost-effective.

Natural rubber is classified into three categories: (i) technically specified rubbers (TSR), where contaminant levels such as ash, nitrogen-containing chemicals, and viscosity are defined; (ii) visually inspected rubbers such as smoked sheets and crepe, which are shipped in 115 kg bales; and (iii) specialty rubbers such as epoxidized natural rubber and powdered rubber. Natural rubber use in radial tires has increased over that used in bias tires because of improved green strength, i.e., strength before vulcanization, tear strength, tire component-to-component adhesion, lower heat generation (hysteresis) under dynamically loaded conditions, and lower tire rolling resistance (Roberts 1988).

Classification of synthetic rubbers is governed by the International Institute of Synthetic Rubber Producers (IISRP). For SBR, polyisoprene, and polybutadiene, a series of numbers have been assigned which define the basic properties of the polymer. For example, the IISRP 1500 series describes the range of commercially available, cold-polymerized emulsion SBR polymers. The series 1200 through to 1249 defines non-oil extended polybutadienes, and 2200 to 2249 is for polyisoprene and copolymers of isoprene.

A series of empirical guidelines can be used in designing a tread compound polymer for a set of tire performance criteria (Nordsiek 1985):

(i) there is a nearly linear drop in abrasion resistance, or tread wear, as polymer glass transition temperature (T_g) increases;

(ii) as T_g increases, wet grip or traction improves;

(iii) inclusion of styrene, and an increase in the styrene level, increases tire wet traction and decreases tire tread wear performance.

Carbon black grades fall into one of typically seven classes: SAF (super abrasion furnace), ISAF (intermediate super abrasion furnace), HAF (high abrasion furnace), FEF (fast extrusion furnace), GPF (general purpose furnace), SRF (semi-reinforcing furnace), and MT (medium thermal). Selection of the appropriate grade of carbon black, further described in ASTM D1765, can, however, lead to significant improvements in performance. For example, the larger surface area carbon blacks (SAF) are required for tread wear. However, as particle surface area increases with a corresponding improvement in tread wear, rolling resistance of the tire can increase with adverse effects on fuel economy. Reduction in carbon black loading will lower tire rolling resistance (see also *Elastomers, Reinforcement of**).

Addition of silica to a compound will improve tear strength, reduce heat build up, and improve compound adhesion in multicomponent products such as tires. Two fundamental properties of silica influence their use in rubber compounds: ultimate particle size and the extent of surface hydration.

Unsaturated elastomers are susceptible to oxidation. Atmospheric ozone will also readily degrade elastomers. For example, tire sidewalls which are subject to a high degree of flexing will show crazing and cracking; such damage is clearly visible in most tires after about four years of operation. To protect an elastomer from oxidation or ozonolysis, three categories of materials are added to the elastomeric formulation, i.e., waxes, antioxidants, and antiozonants (see *Elastomers in Oxygen and Ozone, Degradation of*[#]).

Antioxidants are mainly amine or phenolic derivatives such as 1,2-dihydro-2,2,4-trimethylquinoline (TMQ), and they function by reacting with oxides or broken polymer chain ends caused by reaction with oxygen. They therefore prevent the propagation of oxidative degradation, enabling the retention of the physical properties of the elastomer.

Ozone protection is obtained through the use of two materials, waxes and, typically, paraphenylene diamines. A number of empirical guidelines can be used to develop an antidegradant system for a compound in a tire. Short-term static protection is achieved by use of paraffinic waxes. Microcrystalline waxes provide long-term ozone protection while the finished product is in storage. A critical level of wax bloom is required to form a protective film for static ozone protection. Optimized blends of waxes, paraphenylene diamines such as N-(1,3-dimethylbutyl)-N-phenyl-paraphenylenediamine (6PPD) and antioxidants, such as TMQ, provide long term tire protection under both static and dynamic applications, and over an extended temperature range.

The vulcanization system in a tire compound consists of three components: the activation system, typically zinc oxide and stearic acid; the vulcanizing agent, which is typically sulfur; and accelerators, such as cyclohexylbenzothiazole sulfenamide (CBS). In an idealized tire compound, zinc oxide and stearic acid react to form zinc stearate at tire vulcanization temperatures. The zinc stearate then reacts with the accelerators in the compound to form a "sulfurating complex," which with sulfur creates the required cross-link network. Accelerators can be classified into different categories, depending on the reaction rate and nature of the cross-links they create. Thiurams such as tetramethylthiuram disulfide (TMTD) give predominantly monosulfidic sulfur cross-links. Sulfenamides, when used with high sulfur concentrations, build predominantly polysulfidic cross-links. The structure of the cross-link network can therefore be designed to meet a specific set of service requirements.

A broad range of other materials are used in a compound. For example, titanium dioxide is used for tire white sidewalls; processing oils are used to assist in mixing, extrusion and calendaring of compounds through the factory. Plasticizers such as dioctylphthalate are also used as processing aids. Chemical peptizers such as pentachlorothiophenol are added to the compound during the initial mixing to reduce energy consumption in mixing and provide a more uniformly mixed compound. Resins include extending or processing resins, tackifying resins, which help in handling of components during the tire assembly operation, and curing resins, which are used to improve properties such as tensile strength.

A comprehensive review of materials used in rubber compounding has been compiled by Mark *et al.* (1994), further reference to which is recommended.

5. *Examples of Tire Compounds*

Excluding the weight of the tire reinforcements, a tire's material composition might typically consist of 36% natural rubber, 20% synthetic polymers, 33% carbon black and silica, 6% chemicals, 2% protectants, 2% zinc oxide, and 1% sulfur. Tread compound composition is dependent on the mission of the vehicle for which the tire has been designed. Treads for tires on high-performance cars will use high T_g synthetic polymers, silica, and high levels of oil for traction and vehicle handling. For commercial truck tires, where parameters such as tread wear and fuel economy are more important, high levels of natural rubber and high surface area carbon blacks are used. Wire coat compounds are always composed of natural rubber with a HAF or ISAF grade of carbon black and in some instances silica. Sidewalls are composed of blends of natural rubber and polybutadiene, present in a 1:1 ratio. Carbon black levels are typically 25% of weight of the compound. Other internal components such as shoulder wedges, apex or bead fillers, and base compounds are made primarily of natural rubber, for adhesion, hysteretic properties, and tear strength. Innerliners of radial tires contain chloro- or bromo-isobutene-isoprene rubber (CIIR, BIIR) compounded with HAF, FEF, GPF, or SRF carbon black. Carbon black levels can range from 25% to 50% of the compounded formulation weight, depending on the tire design.

6. *Tire Testing*

Tire testing occurs in three stages: initial laboratory testing; general proving ground testing, where tires are mounted on vehicles which are then tested on specially prepared roads; and commercial fleet tests. Materials tests include studies of vulcanization kinetics, measurements of compound physical properties such as tensile strength and tear strength, and dynamic properties in both shear and tension. Tire laboratory testing includes tire uniformity tests, speed rating determination, rolling resistance, durability assessment, and basic handling characteristics such as cornering coefficient.

Proving grounds consist of high-speed test tracks, gravel roads, wet and dry skid pads, and tethered tracks for testing of farm tractor tires. Testing is conducted under defined conditions such as loads, inflations, speeds, and specific vehicle types. The vehicles have been specified to meet a set of defined conditions such as wheel alignment, vehicle horsepower, wheel base, and axle configuration.

Finally, commercial testing enables the tire materials engineer to obtain an assessment of how compounds and reinforcements will perform under a broad range of service conditions that the end consumer will experience. The result of the testing protocol which tires undergo is a broad range of products to meet the vehicle manufacturers', as well as the end users', needs for optimum performance under a broad variety of service conditions.

Bibliography

American Society for Testing and Materials 2000 Standard classification for carbon blacks used in rubber products. ASTM D1765

Barbin W W, Rodgers M B 1994 The science of rubber compounding. In: Mark J E, Erman B, Eirich F (eds.) *Science And Technology Of Rubber*, Academic Press, Boston, Chap. 9

Bhakuni R S, Chawla S K, Dim D K, Shuttleworth D 1997 *Kirk-Othmer Encyclopedia of Chemical Technology*, 4th edn., **24**.

Clark S K 1982 *Mechanics of Pneumatic Tires*. US Department of Transportation, Washington, DC

Ford T L, Charles F S 1988 *Heavy duty truck tire engineering. 34th Buckendale Lecture SP729*. Society of Automotive Engineers, New York

Kovac F J, Rodgers M B 1994 Tire engineering. In: Mark J E, Erman B, Eirich F (eds.) *Science and Technology of Rubber*. Academic Press, Boston, Chap. 14

Mark J E, Erman B, Eirich F 1994 *Science and Technology of Rubber*. Academic Press, Boston

Nordsiek K H 1985 The "integral rubber" concept – an approach to an ideal tire tread rubber. *Kautsch. + Gummi, Kunstst.* **38**, 178–85

Roberts A D 1988 *Natural Rubber Science and Technology*. Oxford University Press, Oxford

Standards Manual 2000 European Tyre and Rim Technical Organization, Brussels

The Synthetic Rubber Manual, 14th edn. 1999 International Institute of Synthetic Rubber Producers, Houston, TX

The Tire and Rim Association Inc. Year Book. 2000 Tire and Rim Association, Copley, OH

M. B. Rodgers
The Goodyear Tire & Rubber Company, Akron, Ohio, USA

This Page Intentionally Left Blank

S

Sporting Materials: Ski Equipment

The use of sophisticated materials in the design and manufacture of ski boots, bindings, and skis has been a major factor in the advancement of recreational and performance skiing, and "high tech" materials are frequently emphasized in the advertising campaigns promoting new equipment and clothing. Skiing, like tennis, cycling, and golf, has benefited from technological advances that have produced more "user friendly" equipment. In fact, materials development in skis, tennis racquets, golf clubs, and bicycle frames is remarkably similar. This is due in part to the similarities in desired performance characteristics; e.g., lightweight, high strength, durable, responsive, shock absorbing, etc., and the need to tailor these properties to the skill level of the user. From a historical standpoint, Howard Head, an entrepreneur and avid skier, had much to do with transforming both ski and tennis racquet design and the associated materials innovations. Head pioneered the use of metal in ski construction, and much later he introduced the first large-body tennis racquet. He realized the significance of using modern structural materials, and the related design innovations that could be introduced as a result of moving beyond traditional wood structure and manual fabrication techniques.

From the early days of simple wood skis, which Head and others recognized as excessively heavy and difficult to maneuver, the ski industry has graduated to equipment manufactured using more complex, performance-oriented structures. Wood laminates, metal, fiberglass, and plastics were incorporated into ski structures as the materials became available and commercial interest in skiing grew during the 1950s and 1960s. The most important material property contributing to performance was specific stiffness (modulus/density), and glass and carbon-fiber composites had some advantage in this respect over steel and aluminum. Vibration damping, fatigue life, and durability were also improved with composite materials.

The major problem in moving from wood structures to designs incorporating metals and other materials was the integrity of the joints between these dissimilar materials, e.g., wood and steel. Few adhesives were available, and the technology and equipment for assembling and curing complex material combinations was very limited. The development was strictly empirical, and ski testing took place on the slopes; many sets of Head's revolutionary metal skis were subject to premature failure before a successful manufacturing process was developed.

The choice of materials used in modern skis can be linked to specific performance criteria and the trends in ski design, which have paralleled the transition from traditional to contemporary ski technique. Performance is now the single most important issue determining the material of choice, whether the aim is to produce a competitive product for the racer or advanced recreational skier, or an "easy to use" ski for the beginner and intermediate recreational skier. An almost unlimited variety of materials are available and there are markets for even the most expensive products (graphite and specialty composites). Safety and durability are additional factors that influence design decisions, and which are directly linked to the materials selection criteria, specifically in relation to bindings and attachment components, as well as to the structural integrity of the ski. The differences in criteria determining modern ski construction vs. those of the 1950s are compared in Fig. 1. Availability and cost of material, although important, are no longer the limiting factors. Performance and safety are the key issues. Intense competition within the ski industry also promotes the use of "high tech" materials as marketing ploys, thus complicating the picture to some degree. Overall, the modern ski provides a much improved performance, is more user friendly, and is significantly more durable than its predecessor from the 1950s.

As indicated above, the same materials innovations have tended to occur in skis, tennis racquets, mountain bikes, and golf clubs. The common thread in performance is dynamic motion and impact, with a requirement for the athlete to remain in balance and control. Beginners need equipment designed to absorb and insulate them from external impact forces, whereas the experienced athlete can profit from equipment that is more responsive. "Sensitivity" and "feel" are juxtaposed to shock and vibration, and skis are therefore rated according to the skill and performance level of the individual athlete.

For many years skis were characterized as rigid and lively for the racer or advanced athletes, or soft

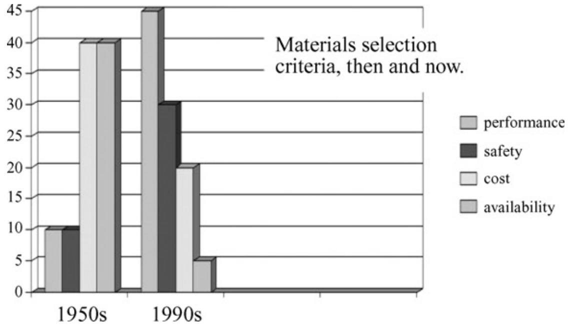


Figure 1
Materials selection criteria of the 1950s and 1990s.

---

* Cross references marked by an asterisk are included in this volume.

[#] Cross references marked by a hash can be found by consulting the Encyclopedia of Materials: Science and Technology.

and damp for the beginner and recreational skiers. The introduction of metal in the early 1950s was a major step leading to the introduction of composite structures engineered to flex and distribute the stresses according to the desired performance criteria. The biggest problem confronting designers like Head was the joining of the dissimilar materials, and the integrity and durability of those joints during use.

Layered structures became quite complex as designers aimed at providing the desired characteristics in the ski, based on the specific use. Specialty alloys, polymer composites, graphite, Kevlar, and ceramic fibers were used to distribute stress, dampen vibration, and determine the manner in which the ski would flex and respond to external forces during use. Materials played an increasingly important role as the construction evolved from a monolithic structure to a multicomponent system incorporating a top layer, side walls, core, edges, and base layer. Some of the more common structures are shown in Fig. 2.

These configurations reflect the requirement for the ski to exhibit significant variable flexibility along its length (to conform to terrain variations), but to have good torsional strength to retain a stable platform for balance and control.

The stress distribution across the surface and the core of the ski is a function of the individual material properties, particularly the yield strength and modulus. The peak stresses can be modulated by the material selection and design considerations. The higher modulus material will carry the bulk of the stress for any given level of strain. This can be shown on a simple stress/strain diagram (Fig. 3).

This can be further illustrated in a simple thin-section composite lay-up under a static load producing an elastic strain (Fig. 4). For the monolith, the stress field is symmetrical with the peak compressive and tension stresses located at the opposing surfaces. The location and distribution of the stress in the layered composite is a function of the material modulus and the position of the layer within the lay-up. For any given amount of (elastic) strain, the higher modulus material will carry a greater portion of the total stress. The peak stress can therefore be concentrated in high-strength, high-modulus material layers within the structure.

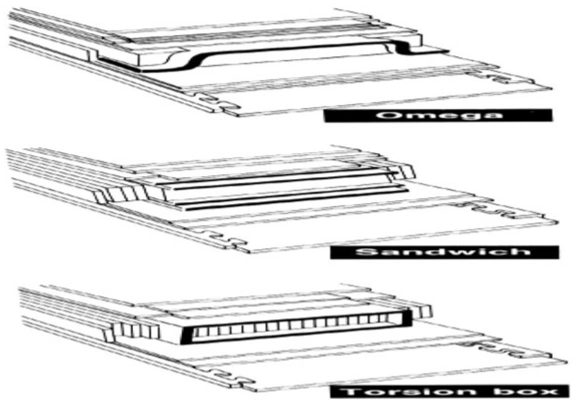


Figure 2
Examples of composite layer structures.

Advanced materials including lightweight metal alloys, carbon fibers, Kevlar, and even piezoelectric ceramic devices have been identified as performance-enhancement factors in recreational skis. The case for using such sophisticated materials is based on the severity of the dynamic loading experienced by the ski in the course of normal operation. This is described in a simplified form as follows.

In operation the ski is subjected to variable dynamic loads at multiple impact points, and is required to flex elastically to conform to extreme terrain variations. The ski must also represent a stable platform on which the skier can balance at speeds which can exceed $50\,\mathrm{km\,h^{-1}}$ for recreational skiers, and significantly higher for racers and advanced athletes. In Fig. 5, the primary forces operating on this platform (the skis) are indicated for one plane only. As long as

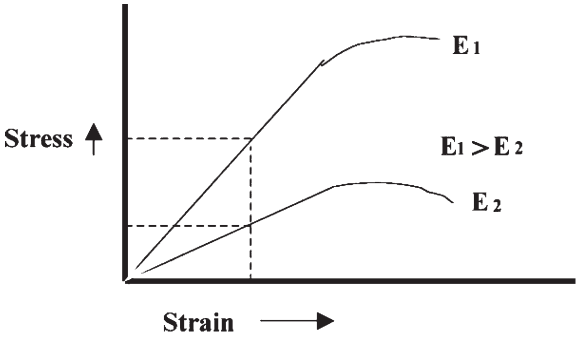


Figure 3
Stress–strain behavior of a typical composite layer material.

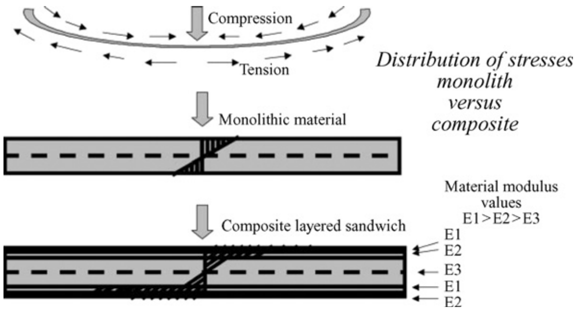

Compression
Tension
Distribution of stresses monolith versus composite
Monolithic material
Material modulus values E1>E2>E3
Composite layered sandwich
E1
E2
E3
E1
E2

Figure 4
Distribution of stresses in monolith vs. composite.

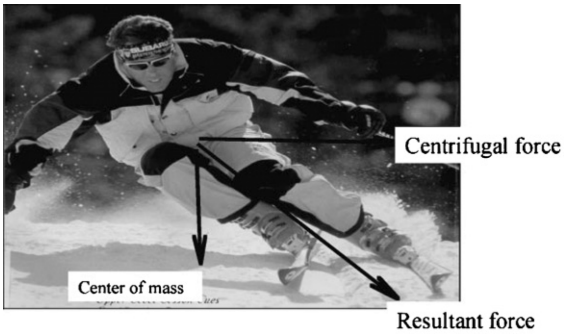


Figure 5
Primary forces operating on skis indicated for one plane only.

the resultant force vector is inside the ski(s)–snow contact points, the skier remains in balance. Any deviation in direction or lateral slippage of the skis requires readjustment by the skier to retain dynamic balance. The strength and rigidity of the ski are therefore key factors determining the performance limits as a function of speed, terrain, weight, and skill level of the skier.

The conflicting requirements (flexibility vs. stability) are addressed by combining dissimilar materials with different functions within the structure, and in some instances employing multifunctional materials to enhance both the efficiency and the performance. The resistance to flexure becomes progressively higher (with amplitude) as the load is transferred to the higher modulus materials in the composite structure. Vibration damping is accomplished with rubber-like absorbing materials in the sandwich layers or in the core. Innovations have favored external damping devices, including mechanical and electromechanical devices such as piezoceramic materials. The latter are based on the dissipation of high-frequency mechanical energy by conversion to electrical energy. Similar concepts are under development for aircraft and automotive structures.

There is a growing awareness of the underlying physics that applies to skiing and the resultant performance characteristics that govern the design and construction details of skis. The move in skis to wider body, "shaped skis" mirrors the similar design trends in tennis racquets, and the enabling technology in both instances has been the materials and related manufacturing development.

There is a vast range of ski equipment that caters for all levels of ability and all types of ski terrain and snow conditions. The choice of construction materials usually accounts for the differences in the performance level. There is no "universal ski" for all abilities and conditions, although most recreational models will provide adequate performance for the average skier under most conditions. Materials will continue to dominate the innovations in ski equipment and ski clothing for the foreseeable future.

H. Casey
Los Alamos National Laboratory, Los Alamos,
New Mexico, USA

Spun (Slurry and Sol–Gel) Ceramic Fibers

Commercial polycrystalline ceramic oxide fibers are produced by spinning and heat treating chemically derived precursors. This chemical process approach is commonly called sol–gel or metal-organic processing. The use of chemical processing allows the preparation of ceramic oxide fibers with high-temperature properties not accessible by traditional approaches, such as spinning of molten glasses. A key characteristic of chemically derived fibers is their ultrafine or even nanoscale microstructure; a grain size below 0.5 μm is optimal for high strength. All commercial ceramic oxide fibers are based on the alumina–silica system since those compositions are particularly useful for high-temperature applications.

Sol–gel fibers are spun in the form of tows or rovings, which typically consist of 400–1000 filaments each having a diameter of 10–15 μm. Fiber tows are flexible and easily handled, which allows them to be woven into fabrics and other complex shapes. The tensile strength of ceramic oxide fibers is typically 1.8–3.0 GPa and their elastic modulus is 150–380 GPa. Major uses for ceramic oxide fibers are as high-temperature thermal and electrical insulation for applications requiring flexibility and light weight, such as sleeves for pipes and electrical cables, high-temperature shielding, belts, blankets, and gaskets. Reinforcement of high-temperature metals and ceramics is an additional area of growing interest.

Advantages of ceramic oxide fibers include stability in air at high temperatures, stability in molten metals such as aluminum, high compressive strength, high elastic modulus, and high electrical resistivity. Disadvantages include low strength compared with SiC and carbon fibers, comparatively high density (3.2–4.0 $\mathrm{g\,cm^{-3}}$), and poor high-temperature creep resistance. The cost of ceramic oxide fibers is higher than that of glass and carbon fibers but lower than that of SiC fibers.

1. Chemical Processing

Chemical or sol–gel processing was first used to produce ceramic oxide fibers in the 1960s (Wainer 1965). A number of chemically derived alumina–silica fibers, notably Saffil (noncontinuous 95% Al_2O_3–5% SiO_2 fibers) and Nextel^{TM} 312 (continuous aluminum borosilicate fibers), were commercialized in the late

1970s (Stacy 1988, Birchall *et al.* 1985, Sowman 1988). Most applications for these fibers are for thermal insulation, where load-bearing capability is not a major concern. More recently, fibers designed specifically for the reinforcement of composites have been produced e.g., DuPont's Fiber FP and Nextel™ 610 and 720 ceramic oxide fibers.

2. *Oxide Fiber Precursor Chemistry*

There are a number of alumina precursors available that are suitable for forming fibers. The aqueous chemistry of aluminum allows for the formation of viscous basic aluminum salt solutions that can be made into fibers by dry spinning. Polymeric aluminoxane precursors have also been used to produce alumina-based fibers. More recently, aluminoxanes that can be melt spun into fibers have been produced. Details of both types of chemical precursor are given below.

The spinning solution or spin dope must be a high-viscosity (10–100 $N\,s\,m^{-2}$ (10 000–100 000 cP)) liquid, stable with respect to both crystallization and precipitation of insoluble salts and organic-containing complexes and to rapid increases in viscosity or gelation resulting from the progression of cross-linking of precursor species. Spin dopes prepared using exclusively colloidal sols (e.g., boehmite, AlOOH) are not suitable for forming continuous fibers because they are shear thinning. Colloidal SiO_2 is, however, commonly used as a component in oxide fiber spin dopes. Other SiO_2 precursors such as partially hydrolyzed alkoxides and polysiloxanes have also been used as a method of introducing SiO_2 into fiber compositions. Other inorganic modifiers to oxide fiber compositions can be readily added using soluble salts (e.g., nitrates). Long-chain organic polymers such as polyethylene oxide are commonly added to improve fiber-forming ability and green strength. Using organic polymers as spinning aids, spin dopes can also be prepared using slurries. Slurry-based fibers (e.g., Fiber FP) have high ceramic content relative to organics, so that shrinkage is reduced, leading to simple heat-treatment. Using slurries, a wide variety of compositions can be produced, limited only by the availability of submicrometer powders (French and Cass 1998). For instance, $BaTO_3$, PZT, YAG, and SiC fibers have been produced using a slurry technique. Sol–gel processes have also been used to prepare a variety of fiber compositions not based on alumina, including ZrO_2, ZrO_2–SiO_2, TiO_2, YAG, and Y_2O_3.

2.1 *Basic Aluminum Salts*

Basic aluminum salts are salts with less than the stoichiometric 3 moles of anion per Al, and have the nominal formula $AlX_n(OH)_{3-n}$, where X is a ligand such as Cl^-, NO_3^-, or CH_3COOH^-. Basic aluminum salts can be formed using a number of methods, including dissolution of aluminum metal in salt solutions, dissolution of aluminum hydroxides in acid, hydrolysis of alkoxides, and neutralization of acidic salt solutions. The polycationic complexes formed are stabilized against further cross-linking and hydrolysis by complexing ligands (Bertsch 1989). Basic aluminum salts used for fibermaking contain 0.5–2.0 moles of anion ligand per mole of aluminum. If the amount of ligand is too low, gelation via cross-linking of polynuclear clusters will occur; high amounts of ligand cause difficulties during fiber heat treatment because excessively large amounts of fugitives must be removed during pyrolysis. Large ligands such as lactic acid or acetyl acetonate can be used to produce spinnable precursors but are not desirable, since ceramic yield on firing is lower than smaller ligands such as chlorine. Aluminum chlorohydrates ($AlCl_x(OH)_y$) are used to prepare a number of commercial fibers, including Fiber FP and PRD-166, Almax™, and Alcen™. Rubilon™ and the Nextel™ series of fibers are produced using aluminum carboxylates.

2.2 *Polymeric Aluminoxane*

Aluminoxanes consist of an Al–O backbone polymer coordinated by chelating ligands such as carboxylates and aceto-acetonates (Glaubitt *et al.* 1994, Yogo and Iwahara 1992). Aluminoxanes are synthesized by the addition of the ligand to aluminum alkyls or alkoxides to partially chelate the aluminum, followed by the addition of 0.5–1moles of water per mole of aluminum to polymerize the aluminoxane. Depending on the ligands used, these precursors can be either melt spun or dry spun. Melt-spun fibers are cured (to make them infusible) by completing hydrolysis by exposure to humid atmospheres. Sumitomo Altex™ fibers are produced using this technique.

Figure 1 shows a schematic of the sol–gel fiber process. Spin dope is extruded through a spinneret to form small-diameter fibers. The fiber becomes rigid as solvent is evaporated. If the viscosity is too low, the extruded liquid spherodizes into droplets; too high, and the fibers become brittle and are easily fractured. After spinning, the green fibers are conveyed into a furnace for heat treatment. Pyrolysis (the conversion of the chemical precursor into a ceramic) typically involves the loss of 50–80% of the initial weight of the green fiber. Volumetric shrinkage can be 80% or higher. Volatile components (including CO, CO_2, H_2O produced by oxidation, plus low molecular weight organics, depending on precursor chemistry) must diffuse out of the fiber during this stage. Thus, pyrolysis must be performed with care to gently decompose the green fiber to the oxide form without forming defects or flaws. Too-rapid evolution of gases can cause pinholes, bubble or void formation, or even bloating, trapping of carbon, phase separation, and cracking. The addition of colloidal or particulate slurries to fiber precursors increases ceramic yield, thereby reducing shrinkage and facilitating pyrolysis.

The goal of heat treatment is to develop a ceramic microstructure with both good strength and good high-temperature properties (i.e., resistance to thermal degradation via grain growth). To achieve high fiber strength at room temperature, a small grain size ($\leqslant 1.0\,\mu$m) is required. The Griffith equation, $\sigma = K_{\rm Ic}/\sqrt{\pi c}$, indicates that for Al_2O_3 ($K_{\rm Ic} = 3$), a strength of 3 GPa can only be achieved if the flaw size c is $< 1.0\,\mu$m. Thus, not only must defect concentration be kept low by careful fiber processing, but grain size should not exceed this limit. During fiber manufacture, extreme care must be taken in all process stages to prevent the formation of defects and flaws that reduce fiber strength.

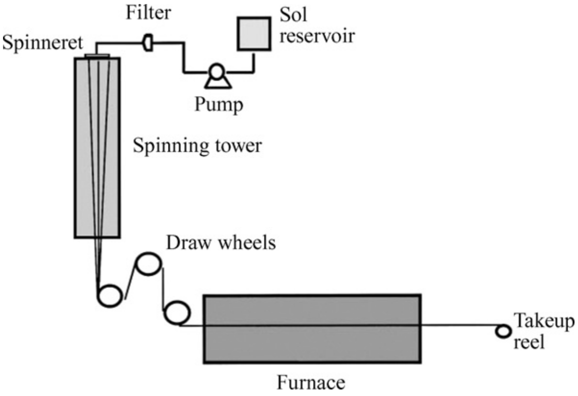


Figure 1
Schematic of sol–gel fiber process.

3. Heat Treatment and Fiber Microstructure

Microstructures in alumina-based fibers must be viewed through an understanding of crystalline transformation sequences in alumina, and the effect of silica on microstructure and transformation kinetics. As shown in Table 1, all commercial oxide ceramic fibers are based on alumina with varying amounts of silica. The stable phases for alumina–silica fibers with $\leqslant 28\%$ Al_2O_3 at all temperatures are α-Al_2O_3 and mullite. However, a series of cubic alumina spinels commonly called transition aluminas form in the temperature range 800–1200 °C during heat treatment of alumina and alumina–silica precursors. The first phase to crystallize is either η-Al_2O_3 (cubic) or γ-Al_2O_3 (tetragonal), which form in the range 800–900 °C. In many systems, δ-Al_2O_3 and Θ-Al_2O_3 are formed with additional heating in the range 1000–1150 °C. These polymorphs are similar to γ-Al_2O_3 but have a higher degree of ordering on the cation lattice. The transformation to α-Al_2O_3 (a hexagonal crystal structure) occurs above 1200 °C; mullite ($3Al_2O_3 \cdot 2SiO_2$) also may crystallize above 1280 °C.

For alumina and alumina–silica fibers, the crystallization of new phases such as mullite and α-Al_2O_3 often result in large grains, causing the fibers to become brittle and weak. A low volumetric nucleation density, which leads to large grain sizes, is typical of both alumina and mullite. For instance, crystallization of α-Al_2O_3 in high-alumina fibers causes the formation of large, porous, spherulitic grains several micrometers in size (McArdle and Messing 1993). These large grains cause the fibers to become brittle and weak. To avoid this problem, most commercial fibers are stabilized as transition alumina by the addition of SiO_2. The addition of SiO_2 to alumina fibers increases the

Table 1
Composition and properties of commercially available oxide fibers.[a]

Fiber	Manufacturer	Composition Al_2O_3–SiO_2	Density	Crystal phase	Tensile strength (GPa)	Elastic modulus (GPa)
Nextel™ 312	3M	62–24 (+14% B_2O_3)	2.7	$9Al_2O_3 \cdot 2B_2O_3$ + am. SiO_2	1.7	150
Nextel™ 440	3M	70–28 (+2% B_2O_3)	3.05	η-Al_2O_3 + am. SiO_2	2.0	190
Nextel™ 550	3M	73–27	3.03	η-Al_2O_3 + am. SiO_2	2.0	193
Nextel™ 610	3M	100–0	3.9	α-Al_2O_3	3.0	373
Nextel™ 720	3M	85–15	3.4	α-Al_2O_3 + mullite	2.1	260
Altex™	Sumitomo	85–15	3.2	γ-Al_2O_3 + am. SiO_2	2.0	193
Alcen™	Nitivy	80–20	3.1	γ-Al_2O_3 + am. SiO_2	1.6	160
Rubilon™	Nichias	68–27 (−5% B_2O_3)	3.0	γ-Al_2O_3 + am. SiO_2	1.8	196
Almax™	Mitsui	100–0	3.6	α-Al_2O_3	1.8	330
Fiber FP	DuPont[b]	100–0	3.9	α-Al_2O_3	1.38	380
PRD-166	DuPont[b]	80 (−20% ZrO_2)	4.2	α-Al_2O_3 + t-ZrO_2	2.3	380

[a] Composition data taken from vendor specifications, precursor composition taken from the patent literature. [b] No longer commercially available.

temperature of the α-Al_2O_3 transformation by as much as several hundred degrees centigrade. This allows the preservation of the submicrometer transition alumina microstructure (and therefore strength and flexibility) for extended periods above 1100 °C. At high silica contents, mullite may crystallize. B_2O_3 is added to some fibers as a grain growth inhibitor; B_2O_3 promotes the formation of a fine-grained mullite structure via $9Al_2O_3 \cdot 2B_2O_3$ intermediate phase (Richards *et al.* 1991). High silica contents will lead to the formation of mullite above 1280 °C, which also can cause weakening of the fibers via grain growth.

Table 1 compares the density, strength, and elastic modulus of commercial ceramic oxide fibers. Commercial fibers can be divided into two classes: (i) those consisting of a mixture of transition alumina and amorphous silica; and (ii) fibers containing crystalline α-Al_2O_3. Alumina–silica fibers have density between 2.7 and 3.2 and elastic modulus between 150 and 200 GPa. α-Al_2O_3 fibers have much higher elastic modulus (380 GPa) as well as higher density (3.6–4.2 $g\,cm^{-3}$). Within each class, density and elastic modulus is inversely proportional to the alumina content: fibers with high alumina content have high elastic modulus and high density. The density and modulus of γ-Al_2O_3 is in the range 3.2–3.6 $g\,cm^{-3}$ and 220 GPa respectively; the addition of SiO_2 at 2.2 $g\,cm^{-3}$ and 70 GPa will reduce both. The addition of B_2O_3 will also reduce density and modulus. The effect of silica content on strength is less clear, since strength is related to flaw size and distribution as well as composition. Most fibers have tensile strength in the range of 1.7–2.1 GPa. An exception is Nextel 610, which has a strength of 3.0 GPa.

Fibers containing crystalline α-Al_2O_3 have much higher chemical stability and elastic modulus than fibers containing transition aluminas. They are also not prone to shrinkage at high temperatures caused by crystallization and sintering, and can have higher creep resistance under load at high temperatures. For these reasons, they are used for the reinforcement of metal and ceramic composites; however, achieving fine grain size in α-Al_2O_3 fibers is not straightforward. α-Al_2O_3 has an intrinsically low volumetric nucleation density, which leads to large grain size and low strength. The addition of seeds or nucleation agents to promote high nucleation density during crystallization is essential if the fine grain sizes required for high strength (0.5 μm) are to be achieved. Seeding with submicrometer α-Al_2O_3 particles (e.g., Fiber FP) or Fe_2O_3-based nucleation agents (e.g., Nextel 610) have been used to produce fine-grained, high-density α-Al_2O_3.

Ceramic oxide fibers have limited creep resistance at high temperatures compared with SiC fibers; however, progress in fiber processing during the 1990s has enabled the preparation of fibers with significantly improved high-temperature creep properties. Fibers with increased amounts of crystalline phases such as α-Al_2O_3 and mullite, and lesser amounts of amorphous phases such as SiO_2, have reduced deformation at high temperatures. For instance, Nextel 720 fiber, which consists of a mixture of α-Al_2O_3 and mullite, has demonstrated considerably greater creep resistance above 1000 °C than either alumina–silica or α-Al_2O_3 fibers. This has been attributed to the presence of mullite, which is a more creep resistant phase than alumina, plus larger grain size. YAG fibers that exhibit even better creep resistance have been developed at several laboratories. Such YAG fibers have the potential for long-term composite reinforcement at 1200 °C and above.

Bibliography

Bertsch P M 1989 Aqueous polynuclear aluminum species. In: Sposito G (ed.) *Environmental Chemistry of Aluminum.* CRC Press, Boca Raton, FL, pp. 88–111

Birchall J D, Bradbury J A A, Dinwoodie J 1985 Alumina fibres: preparation, properties and applications. In: Watt W, Perov B (eds.) *Handbook of Composites.* Elsevier Science, Oxford, UK, Vol. 1, pp. 115–54

French J D, Cass R B 1998 Developing innovative ceramic fibers. *Am. Ceram. Soc. Bull.* **77**, 61–5

Glaubitt W, Sporn D, Jahn R 1994 A new way to spinnable sols derived from modified aluminum alkoxides. *J. Sol-Gel Sci. Technol.* **22**, 525–8

McArdle J L, Messing G 1993 Transformation, microstructure development, and densification in α-Fe_2O_3-seeded boehmite-derived alumina. *J. Am. Ceram. Soc.* **76**, 214–22

Richards E A, Goodbrake C J, Sowman H G 1991 Reactions and microstructure development in mullite fibers. *J. Am. Ceram. Soc.* **74**, 2404–9

Sowman H G 1988 Alumina-boria-silica ceramic fibers from the sol–gel process. In: Klein L (ed.) *Sol–Gel Technology for Thin Films, Fiber Preforms, Electronics and Spec. Shapes.* Noyes, Park Ridge, NJ, pp. 162–83

Stacey M H 1988 Developments in continuous alumina-based fibres. *Br. Ceram. Trans. J.* **87**, 168–72

Wainer E, Raynes B C, Cunningham L 1965 Liquid polymers, solid articles made therefrom and methods of preparing same. US Pat. No. 3,180,741

Yogo T, Iwahara H 1992 Synthesis of α-alumina fibre from modified aluminum alkoxide precursor. *J. Mater. Sci.* **27**, 1499–504

D. M. Wilson
3M Co., St. Paul, Minnesota, USA

Superconducting Wires and Cables: Materials and Processing

1. Introduction

The phenomenon of superconductivity was first observed in 1912 in the laboratory of Heike Kamerlingh Onnes, at the University of Leiden (Holland). At low temperatures certain materials (~ 800 superconducting compounds have been recently tabulated by Poole and Farach 2000) suddenly lose their resistance to the flow of electricity. The phenomenon remained a laboratory curiosity until 1954 when G. B. Yntema at the University of Illinois made the first successful superconducting magnet. At that point, magnet performance fell well below that predicted by the properties of the superconductor. The slow progress can be partially attributed to (i) the difficulty in obtaining the extremely low temperatures required and (ii) the lack of understanding of how to create electrically and thermally stable wires from superconductors. The primary limitation, however, was that the superconducting phenomenon was not only limited to low temperatures but also to a restricted range of electrical current density and magnetic field and that the range of these properties was particularly limited in early superconductors. The maximum values of temperature, electrical current, and magnetic field are interdependent and, when plotted in three axes, form a "critical surface" (Fig. 1). Those early superconductors, such as Pb, In, and Hg are classified as type I superconductors. These are superconductors in which magnetic flux is excluded from their bulk and the critical current density is limited to a surface layer of approximately one-tenth of a micrometer. The maximum fields in which these superconductors can operate are usually less that 0.1 T (similar to the flux-density between the poles of a horseshoe permanent magnet). These limitations make type-I superconductors impractical for wire and cable applications. All "technical superconductors," i.e., superconductors that can be readily fabricated into wires and cables for high-current applications, are type-II superconductors. In type-II superconductors, magnetic flux penetrates the bulk of the superconductor to form individual flux quanta (see *Electrodynamics of Superconductors: Flux Properties*[#]). In the most widely used technical superconductors, Nb–Ti and Nb_3Sn, the upper critical fields are 13 T and 27 T, respectively, and current densities $>10^9\,A\,m^{-2}$ can be carried (compared to $\sim 10^7\,A\,m^{-2}$ for domestic Cu wire). In Fig. 2, the critical current densities at 4.2 K are compared for strands or tapes that are fabricated in lengths of more than 100 m. The temperature range is still limited (the critical temperatures for Nb–Ti and Nb_3Sn are 10 K and 18 K, respectively), but for many applications the cost of insulation and refrigeration is more than offset by the reduced energy costs resulting from nonresistive current flow. Because the amount of energy lost in current flow through a superconducting magnet is extremely low, it can be operated without a power supply once the magnet has been charged. This mode of operation, termed "persistent" is very desirable for MRI and NMR applications because the current flow (and consequently the field) is very stable. Liquid He is now readily available and provides cooling to 4.2 K at atmospheric pressure and as low as 1.8 K at reduced pressure. Recent advances in insulation and

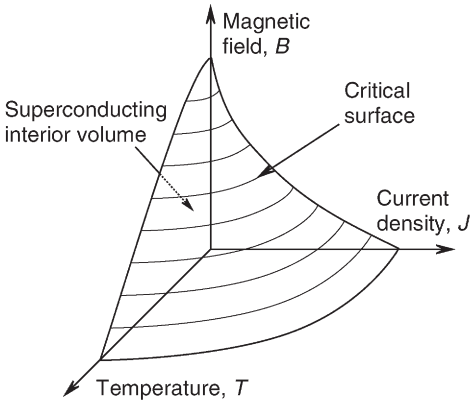


Figure 1
Schematic illustration of the critical surface for a type-II superconductor. The applied field (*B*), the temperature (*T*), and the current density (*J*) must be maintained below the critical surface in order to retain superconductivity.

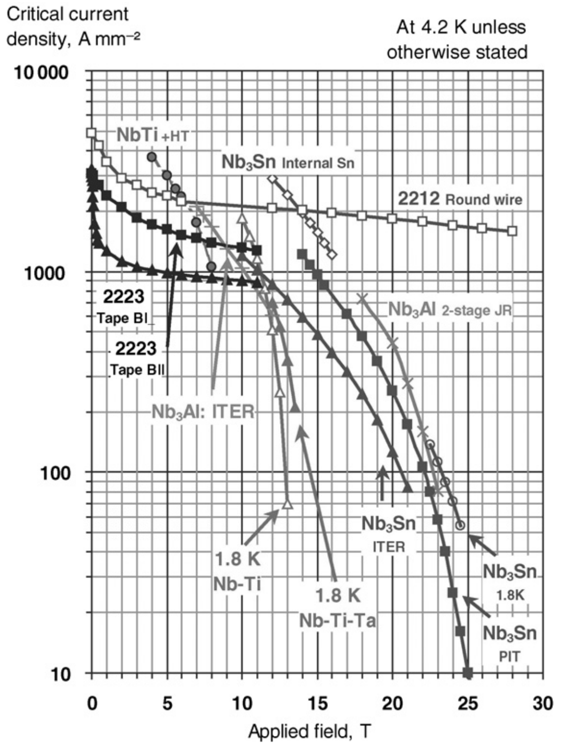


Figure 2
Comparison of critical current densities (at 4.2 K unless otherwise stated) available in strands and tapes of more than 100 m.

refrigeration technology have also made the application of superconductors more widespread. The discovery of high-temperature superconductivity (HTS) superconductors (ceramic superconductors with critical temperatures above the boiling point of liquid nitrogen (77 K) in 1986 (see *High-temperature Superconductors: Thin Films and Multilayers*[#])) dramatically increased interest in the application of superconductivity. The discovery of superconductivity at 39 K in MgB_2 in 2001 further extended the temperature range for intermetallic superconductors.

Superconductors must be formed into composite wires in order for them to be used in high-current applications and these wires may also be cabled. The wide variety of mechanical properties of superconductors: the range from strong ductile alloys like Nb–Ti to brittle ceramics such as HTS superconductors means that the wire processing routes vary greatly. Of the HTS superconductors only one, Bi-2212, is manufactured in round-wire form and HTS superconductors are more commonly manufactured as multifilamentary tapes. The brittle superconductors cannot undergo wire drawing and are consequently mechanically processed in the form of ductile precursors or fine powders that are reacted to form the final superconductor. For Nb–Ti, the alloy can be cold-worked to a remarkable extent but heat treatments are also applied during processing in order to introduce a second phase into the microstructure that will ultimately pin the fluxoids in the final wire. Because of these variations, the types of processing will be introduced separately and by superconducting material. There are, however, key components that are common to all superconducting strands.

2. General Strand Requirements

2.1 Matrix and Filaments

Superconducting wires are almost always of a multifilamentary design in which individual continuous filaments ($< 30\,\mu m$ in diameter) are embedded in a high-conductivity matrix. When conditions in a localized region of a superconductor exceed the critical surface conditions, that area will start to conduct resistively and produce local heating. The resulting heat and current transfer will start a cascade effect and result in a rapid transition of the whole superconductor to the normal state. The ability to transfer heat and current away from localized regions by subdividing the filaments into a high-conductivity matrix was a key step towards producing practical superconductors. The primary components of cryogenic stabilization were not formulated until the mid-1960s (Stekley and Zar 1965). The most desirable matrices are Cu or Al, but Ag is required for the processing of HTS superconductors where oxygen diffusion through the matrix is required during heat treatment. External stabilizer is often added after heat treatment; this is particularly the case for Al, which has thermal and density advantages over Cu, but is difficult to fabricate as the matrix of a composite wire. In addition to stability requirements, fine filaments also serve the purpose of limiting hysteretic loss, not only for a.c. applications but also for d.c. applications. The filament diameter requirements for a.c. applications are quite severe with desired filament diameters of less than 1 μm. For a.c. applications, a Cu–Ni matrix is often used to reduce proximity coupling of the filaments.

2.2 Filament Uniformity

As the critical current per unit cross-sectional area is limited by the material properties, the amount of current that can be carried by an individual filament is reduced if the filament is locally necked to a smaller diameter. Current may be locally transferred to adjacent filaments, but there is a broadening of the critical current transition and a reduction in time that the strand may be operated in the persistent mode (crucial for MRI application). With final filament diameters of less than 20 μm often required, this imposes a very high level of quality control on the processing of the strands. The degree of filament uniformity is reflected in the sharpness of the superconducting transition with increasing current. In the early stages of the resistive transition, the curve of voltage, V, against current, I, can be described by the power law:

$$V \propto I^n \tag{1}$$

The value of n (the transition index) decreases with field; the more uniform the filament cross-sectional areas, the more linear the decrease in n-value and the higher the initial low field value (Warnes and Larbalestier 1986). An electric field, E, is generated along the superconductor carrying a current density, J, that is related to the critical current density of the superconductor J_c at field E_c by the transition index, n:

$$E = E_c(J/J_c)^n \tag{2}$$

For composites based on Nb–Ti strands, the value of n (5 T) can be 50–100, in Nb_3Sn for NMR application the values are in the range of 40–80. The high n-values in Nb–Ti strands correspond to variations in filament cross-sectional area of less than 2% (coefficient of variation) (Lee and Larbalestier 1993). In HTS superconductors the range of available n-values is presently only 10–20 and persistent mode operation is not feasible under these conditions.

2.3 Critical Current Density and Flux Pinning

Pure annealed superconductors do not carry very much current because once current passes through the superconductor, a Lorentz force is induced on the flux lines. When they move, the flux line lattice (FLL) dissipates energy and the superconductor eventually goes "normal" (see *Electrodynamics of Superconductors: Flux Properties*$^{\#}$: *Superconducting Materials: Irradiation Effects*$^{\#}$). The flux lines, however, can be held in place by defects in the superconductor, a process termed "flux-pinning." In A15 based superconductors (such as Nb_3Sn), increasing the grain boundary density has been shown to increase critical current density (Scanlan *et al.* 1975, Ochiai *et al.* 1986). In Nb–Ti a linear increase in critical current density is observed with vol.% of α-Ti precipitate (Lee *et al.* 1990). Consequently the introduction of pinning defects and the refinement of grain size are key elements to the success of technical superconductors.

2.4 Twisting

Twisting of the strand about its drawing axis is typically required to reduce flux-jump instability caused by varying external fields, and to reduce eddy-current losses. The twisting is ideally applied just before a multifilamentary strand has reached final size, so that the twist can be locked into place by a further die-pass. The higher the expected rate of change of field, the tighter the required twist pitch. For the relatively steady-state dipole magnets for the Superconducting Supercollider, ~80 rotations along the drawing axis per meter were required, for a.c. application with a similarly sized strand the number of twists per meter might be 300. Typically, twist pitches cannot be higher than 8 times the wire diameter, and normally a twist pitch of 10–20 times the diameter is used.

2.5 Final Shaping and Cladding

The final cross-section of the strand can be controlled by die shape and size or by using independently adjusted rollers operating along the strand surface. In this way, square or rectangular cross-section filaments can be produced. Alternatively, the strand can by inserted into a channel in a larger form, most commonly a rectangular aluminum external stabilizer.

2.6 Cabling

Individual strands can be cabled or braided together to form a conductor with a higher current carrying capacity. The most common design for Nb–Ti magnets is the Rutherford cable, which consists of a fully transposed, flat cable. Using this approach, high-aspect-ratio cables can be produced with as many as 46 strands (Scanlan *et al.* 1997). Strands may also be drawn into a larger conduit or combined with external stabilizers. A cross-section of a cable-in-conduit-conductor (CICC) is shown in Fig 3. As in the single strand-in-channel aluminum process, multiple strands have been applied to surface channels in circular cross-section aluminum rods.

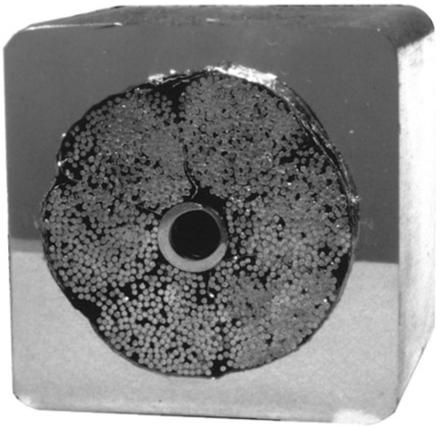


Figure 3
CICC: this 51 mm diameter Incoloy® jacketed cable contains 6 sub-cables; each containing 180 Nb_3Sn based superconducting strands.

2.7 Length-Dependent Parameters

Both commercial Nb–Ti and Nb_3Sn strands are made in lengths of over a kilometer and variations can occur over the length of the strand, due to differences in mechanical properties of the components and bonding of the monofilaments that occurred during mechanical reduction. In Nb–Ti variations in critical current, I_c, along the length of strands can be directly correlated with variations in Cu/Nb–Ti ratio along the length as shown by Kanithi *et al.* (1990). In their study of strands developed for the Superconducting Supercollider project, they showed that the variation in I_c (due to variations in the intrinsic superconducting properties, the strand diameter and Cu/Nb–Ti ratio as well as measurement error) was more than twice the variation seen in J_c (due to just the intrinsic superconducting properties and measurement error). As HTS-based strands are still in their production infancy, a different phenomenon is observed: as production length is increased, the overall I_c declines because it is very difficult to maintain the ideal conditions (grain alignment, heat treatment temperature, low porosity) necessary for good HTS strand over long lengths.

3. Superconducting Strand and Cable Manufacture by Material

3.1 Nb–Ti

(*a*) *Alloy Composition.* Nb–Ti has become the dominant commercial superconductor, because it can be economically manufactured in a strong and ductile form with both high current density and high resistive transition index. The starting point for fabrication of Nb–Ti strand is an alloy of between 46.5 wt.% Ti and 50 wt.% Ti. Decreasing Ti over this range increases both the critical temperature, T_c, and the upper critical field, H_{c2} (which at 4.2 K has a maximum value of 11.5 T at 44 wt.% Ti). The increase in T_c and H_{c2}, however, is offset by a decrease in the amount of Ti available for precipitation of the α-Ti phase desired for flux-pinning and high critical density. The composition of the Nb–Ti also affects the precipitate morphology: if the α-Ti is formed at grain boundary intersections, it tends to be uniform in size and distribution and does not significantly affect strand ductility. Higher Ti compositions, however, also produce fine distributions of Widmanstätten-type precipitation in the grain interiors resulting in both a nonuniform microstructure and a significant hardening of the Nb–Ti. As all these factors are sensitive to composition, it is very important to start with a homogeneous alloy. With final filament diameters well below 0.05 mm, the alloy must also be free of unmelted Nb or nonductile inclusions that may result in filament and ultimately strand breakage. Although the large difference in the melting points of Ti and Nb makes this a difficult alloy to manufacture, homogeneity levels in commercial Nb–47 wt.% Ti alloys are now typically much better than ± 1.5 wt.% Ti. Impurity content has been traditionally kept low and consistent, but recent work has shown that increasing the allowable content of iron can considerably refine the precipitate size without adversely affecting the precipitate volume (Smathers *et al.* 1996).

(*b*) *Extrusion and Billet Assembly*. The (as-cast) Nb–Ti ingot is hot forged to the required diameter and then annealed in the single phase β region (~ 2 h at 870 °C). For fine filament Cu matrix composites, the Nb–Ti ingot is then sealed inside a high-purity Cu extrusion can, that subsequently (warm-extruded) forms a (well-bonded) monofilament. Monofilament production billets ready for extrusion are shown in Fig. 4. Billet assembly is performed in a clean environment, in order to avoid introducing particles that would not co-deform to the sub-0.05 mm filament size. Where high critical current densities and fine filaments are desired, an Nb diffusion barrier is wrapped around the Nb–Ti rods in order to inhibit the formation of brittle Cu–Ti intermetallics during heat treatment (Wilson 1972). The monofilament is then reduced in diameter by wire drawing, typically finishing with a hexagonal die to produce filaments that can be efficiently (restacked). The thickness of Cu around the Nb–Ti at this stage is important because it determines the spacing between Nb–Ti filaments when Cu/Nb–Ti monofilaments are stacked together. For optimum mechanical stability in Nb–Ti/Cu multifilamentary composites, a filament spacing to filament diameter (s/d) ratio of 0.15–0.20 is ideal (Gregory *et al.* 1987). Usually additional stabilizer is required and that may be added by increasing the wall thickness of the extrusion can or adding a central core of stabilizer. For d.c. magnet application, it has been established by Ghosh *et al.* (1987) that a minimum Cu thickness of 0.4–0.5 μm is required to reduce the magnetization associated with proximity filament coupling. For s/d ratios of 0.15–0.20, the minimum Cu thickness requirement limits the minimum filament diameter to ~ 3 μm when using a pure Cu matrix. If smaller filaments are required, Ni or Mn can be added to the Cu between the filaments (Hlasnik *et al.* 1985, Cave *et al.* 1989, Kreilick *et al.* 1988).

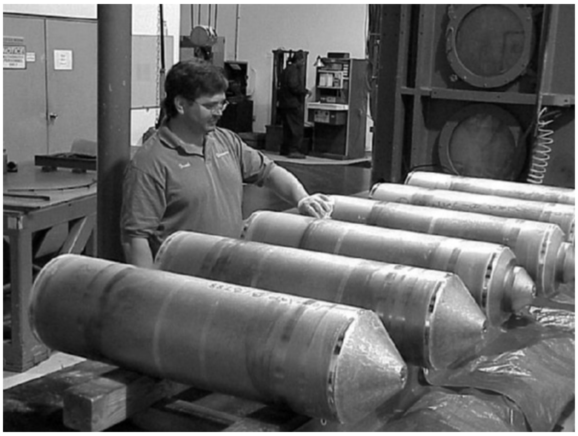


Figure 4
Monofilament Cu-clad Nb–Ti extrusion billets being inspected after welding (image courtesy of Hem Kanithi, Outokumpu Advanced Superconductors Inc.).

The monofilament is then restacked inside another Cu extrusion can to form the final multifilamentary composite. The size of this extrusion can be up to 330 mm in diameter with a weight of 400 kg. As many as 4000 filaments may be assembled in one extrusion. Alternatively, large numbers of fine filaments can be produced by restacking a previously extruded multifilamentary billet. Using this approach, strands with up to 40 000 filaments have been manufactured. Where a small number of large filaments are desired (typically for low-field MRI application), Nb–Ti rods can be directly inserted into gun-drilled holes in a solid high-purity Cu extrusion billet. Warm extrusion helps bond the composite, as much of the subsequent processing involves cold work by wire drawing, and the resulting strain introduces into the Nb–Ti an essential element in developing the high critical current microstructure.

(*c*) *Precipitation Heat Treatment*. Such a treatment qis applied after sufficient cold work has been carried out following the multifilamentary extrusion. If too little cold work is applied before heat treatment, Widmanstätten-type precipitation will occur, causing increased hardness (reducing drawability) and producing a non-optimum precipitate size and distribution. Lee *et al.* (1989) found that the amount of (prestrain) required increases linearly with wt.% Ti. Consequently, a higher prestrain is required to suppress Widmanstätten-type precipitation in a chemically inhomogeneous alloy than would be required in a homogeneous alloy. For Nb–47 wt.% Ti, the required prestrain is ~5 but this rises steeply to 9 for a 55 wt.% Ti alloy. Heat treatments are typically at 375–420 °C, ranging from 20 h to 80 h in duration. Multiple heat treatment and drawing cycles are required to produce high critical current densities (Cheng-ren *et al.* 1983). Whereas an initial heat treatment will only yield ~10 vol.% precipitate, by applying additional cold work strain, the effect of the slow diffusion rate at these temperatures is again overcome and more precipitate is produced. Second and third heat treatments typically yield 15 vol.% and 20 vol.% of α-Ti precipitate respectively. The strain between heat treatments is 0.8–1.5 and at least three heat treatments are usually applied. A linear relationship between the optimized critical current density and the volume of precipitate Nb–47 wt.% Ti alloys has been established by Lee *et al.* (1990) 20% volume of α-Ti precipitate is sufficient to produce a critical current density of >3000 A mm^{-2} at 5 T, 4.2 K. Increasing Ti content increases the precipitation rate and amount, but higher pre-strains are required if good ductility and long piece length is to be maintained. Alternatively, the pinning sites can be introduced mechanically by hand-assembling the desired microstructure at a large-size, in the same way as the multifilamentary billets are assembled, and then reducing the dimension to match the fluxoid spacing by multiple restacking and drawing cycles. This approach is called APC for artificial pinning center. With APC strand, no additional heat treatment is necessary and a greater flexibility is possible in the microstructural components. Despite superior low field (<5 T) performance, the APC approach has not reached commercialization because of the costs associated with the multiple restacking approach.

(*d*) *Final Wire Drawing*. After final heat treatment the α-Ti precipitates are roughly equiaxed in transverse cross-section, 100–200 nm in diameter with an extension along the wire axis giving an aspect ratio of 5–20. The extensive final wire drawing (an engineering true strain of 5), however, produces a plain strain condition in the b.c.c. β-Nb–Ti, which is supported by (inter-curling) of the Nb–Ti grains. This results in distortion of the α-Ti precipitates into densely folded sheets during final wire drawing (Fig. 5(a)). The folding process rapidly decreases the precipitate thickness and spacing with a dependence of $d^{1.6}$ (where d is the strand diameter) and increases the precipitate length per area with a dependence of $d^{-1.6}$, as measured by Meingast *et al.* (1989). The critical current density increases as the microstructure is refined until it reaches a peak, after which there is a steady decline. The peak in critical current density for a monofilament or a multifilamentary strand with uniform filaments occurs at a final strain of ~5. If the filaments are nonuniform in cross-section (sausaged), the peak occurs earlier and at a lower critical current density. A strand that has a premature (and lowered) peak in critical current density during final drawing is described as "extrinsically limited" because it has not attained the intrinsic critical current density of the microstructure. The most common source of extrinsic limitation is sausaging of the filaments, due to intermetallic formation or lack of bonding between the

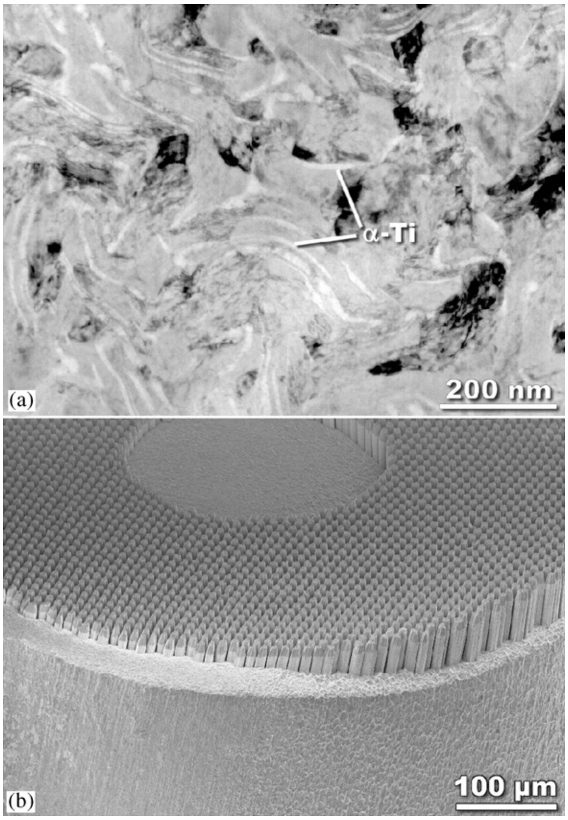


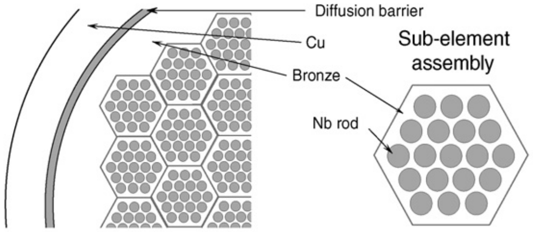
Figure 5
(a) Transmission electron microscope image of high critical current density Nb–Ti microstructure with folded sheets of α-Ti precipitate. (b) Etched cross-section of Nb–Ti-based superconducting strand fabricated by IGC-AS (now Outokumpu Advanced Superconductors) for the Interaction Region Quadrapole Magnets of the Large Hadron Collider at CERN.

components of the composite. A high-performance strand (Fig. 5(b)) has sausaging reduced to a very low level—a coefficient of variation for the filament cross-section of $\sim 2\%$. With tight quality control, uniform properties and piece lengths exceeding 10 km should be expected.

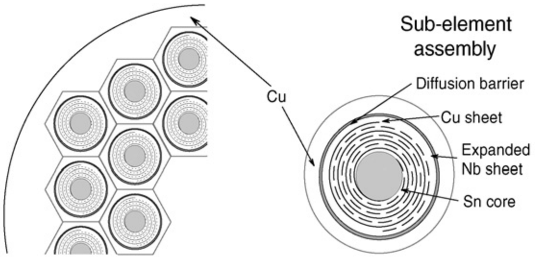
3.2 Nb_3Sn

Nb_3Sn is the second most widely used superconductor and has a higher upper critical temperature (18.2 K) than Nb–Ti. It is one of several A15 structure compounds that are superconducting. Other A15 superconductors of interest include Nb_3Al (T_c of 19.1 K) and Nb_3Ge (T_c of 23.2). Nb_3Al has been manufactured in long lengths and has a better strain tolerance than Nb_3Sn and better performance at a very high field (> 22 T) making it of importance to high-field NMR inserts. Nb_3Al is much more complex and expensive to fabricate than Nb_3Sn because of the mechanical differences between Nb and Al and because the optimum Nb_3Al composition for a A15 structure, can only be formed by high temperature processing (> 1000 °C). Nb_3Sn, on the other hand, is widely manufactured and has shown sufficient strain tolerance in practice to be used successfully in very large solenoids such as the ITER CS model coil which stored 640 MJ of energy (Kato *et al.* 2001) and high-field (> 21 T) inserts for NMR application (Kiyoshi *et al.* 2001). Because A15 materials are brittle, all production routes use ductile precursors that are heat treated to form the A15 at final size by solid-state diffusion. At an early stage it was found that the presence of Cu greatly aids the formation of Nb_3Sn at temperatures below 800 °C. It has also been found that adding Ti (0.7–1 wt.%) or Ta (7.5 wt.%) improves the critical current density at high fields (> 12 T).

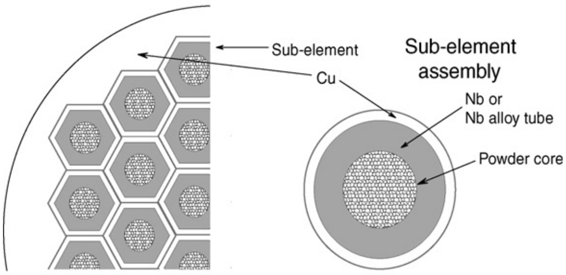
The three basic production routes currently used by commercial Nb_3Sn vendors—the bronze process, the internal Sn, and the powder-in-tube (PIT) fabrication techniques—are illustrated in Fig. 6. The internal Sn process may use a jelly roll of Cu and Nb sheets, as illustrated in Fig. 6, or may use Nb (or Nb alloy) rods in a Cu matrix as shown in Fig. 7.



Bronze process



Modified jelly roll



Powder-in-tube

Figure 6
Schematic illustration comparing the bronze process, modified jelly roll (internal Sn), and PIT fabrication techniques for Nb_3Sn-based superconducting strand.

(*a*) *Bronze Process.* For MRI application the most common route is the "bronze process" in which Nb alloy rods are first clad in high-purity ductile Cu–Sn bronze and then assembled inside another Cu–Sn tube in a way similar to the production of Nb–Ti filaments in Cu for Nb–Ti strand. As with Nb–Ti alloy, multiple restacks can be used to increase the filament number. Under heat treatment (typically 650–725 °C) the Sn from the bronze reacts with the Nb to form the superconducting A15 compound. As the Cu–Sn alloy does not act as a stabilizer, additional high-purity Cu must be added outside, and/or inside the bronze–Nb filament composite, with a Ta and/or Nb diffusion barrier to eliminate Sn contamination of the stabilizer. An example of a bronze-process strand is shown in Fig. 8. Vanadium has also been used as a barrier but results in a reduced critical current density. This process can yield a strand with a very high uniformity in filament size and distribution, which makes it an ideal product for high-field NMR and other applications where a minimum of hysteresis loss is desired. The bronze, however, quickly work hardens during wire drawing and must be annealed frequently in order to maintain strand ductility, this extra processing step significantly increasing the cost of the strand. The Sn content in the composite is also limited by the amount of Sn (< 15 wt.% Sn) that can be added to bronze while maintaining ductility. The impact of insufficient Sn is most

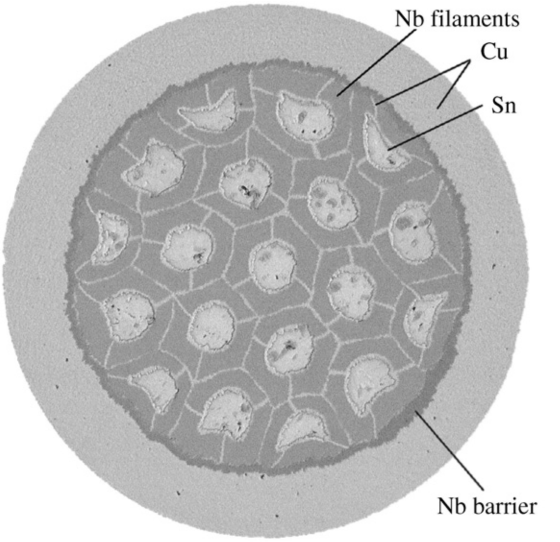


Figure 7
Transverse cross-section of a rod-process internal Sn Nb_3Sn strand prior to diffusion heat treatment. This strand was designed for high critical current density by IGC-AS (now Outokumpu Advanced Superconductors).

noticeable on critical current density. Recent developments in the bronze process have centered on the use of specially processed high Sn bronzes (up to 16 wt.% Sn) that have good ductility and high homogeneity (Sakamoto *et al.* 2000) (Fig. 9).

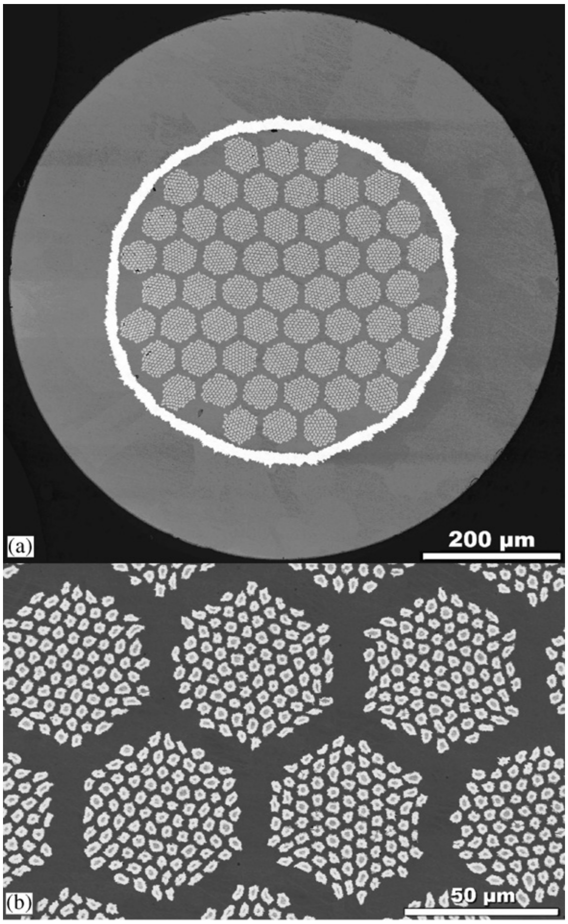


Figure 8
Bronze process Nb_3Sn strand fabricated by Vacuumschmelze for ITER shown in (a) full cross-section and (b) in filament detail. The filaments have a central unreacted core of Nb.

(*b*) *Internal Sn Process*. Higher Sn content and higher critical current density can be achieved by keeping the Sn separate from the Cu during strand fabrication. This is the basis for the internal Sn process which places an Sn core inside an Nb–filament/Cu matrix. The Sn core can be alloyed as a carrier for the Nb_3Sn allowing-element or to improve the mechanical properties of the soft Sn. Strands produced by this process have achieved the highest critical current densities for Nb_3Sn (in excess of 2400 A mm^{-2} in the non-stabilizer package). The high critical current densities are achieved, in part, by reducing the amount of Cu in the Cu/filament/Sn package and the resulting reduction in filament coupling results in an undesirable increase in filament coupling. There are two distinct sub-categories of internal Sn process strand based on the form of the Nb filaments. In the "rod-in-tube" (RIT) process internal Sn–Nb filaments are assembled by inserting rods of Nb into pre-drilled holes in a Cu billet. The second internal Sn method is termed modified jelly-roll (MJR) and used an expanded metal sheet of Nb co-wrapped with a Cu sheet to create the individual Nb filaments in a Cu matrix. Both of these techniques achieve very high current densities, but limitations to the sizes of the billets (because of the restricted depth of the gun-drilled holes for rod-process and size limitations to deforming the MJR stacks) means that billet sizes are typically only 20 kg and processing costs are subsequently much higher than for Sn. As in the bronze process, the individual Cu–Nb–Sn sub-elements are clad in a diffusion barrier (Ta or Nb or TaNb alloy) to prevent poisioning of the Cu stabilizer by Sn diffusion. Sn diffuses rapidly through Cu at the temperatures used for the A15 forming heat treatment (up to 700 °C) and barrier integrity is consequently of great concern. The heat treatment of the strands to form the A15 phase is preceded by multiple lower-temperature steps designed to combine the Sn with

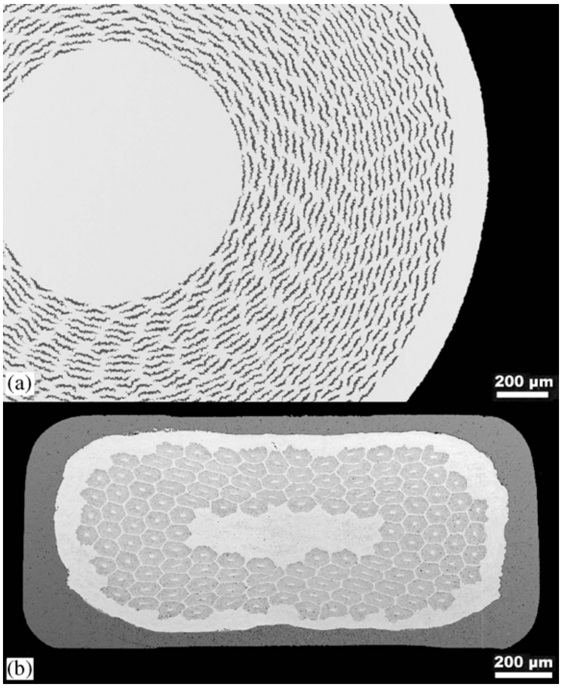


Figure 9
(a) Precursor jelly-roll of Al and Nb fabricated by Oxford Advanced Superconductors. (b) Multifilamentary Nb_3Al tape fabricated by the National Institute for Materials Science in Japan from a rapidly heated and cooled composite of jelly-roll precursors (fabricated by Hitachi Wire and Cable).

the Cu without Sn melting. Both the RIT and MJR technique can also be used to fabricate precursor strand for Nb_3Al production.

(*c*) *PIT Process*. Nb_3Sn may also be formed from higher Sn content Nb–Sn phases (Nb_6Sn_5 and $NbSn_2$) by inserting the intermetallics in powder form into tubes of Nb and reacting at 650–700 °C. The presence of Cu is also required to allow the formation of the A15 phases in the 650–700 °C temperature range in which fine grain size is maintained (Lefranc and Muller 1976). Elemental Sn is also added to improve the Sn-to-Nb ratio. No low-temperature heat treatment is required for these strands and the A15 formation heat treatment takes a relatively short time, 50–75 h (180–270 ks). This process uses the best combination of high critical current densities ($>2200\,\text{A}\,\text{mm}^{-2}$ at 12 T and 4.2 K in the nonstabilizer volume) and low effective filament diameter (as low as 30 μm in high critical current density strand). The process is used to manufacture small quantities for demanding applications (Lindenhovius *et al.* 2000) and has not received the cost benefits of commercial scale-up. The PIT technique can also be used for MgB_2, Bi-2223, and YBCO, but porosity is an issue for these conductors when fabricated by this technique.

3.3 Nb_3Al

Nb_3Al has not achieved commercial strand status, but it can be described as a technical superconductor as it has been fabricated into high performance and uniform kilometer lengths that have in turn been fabricated into full-scale devices. At magnetic fields above 20 T, critical current densities can be obtained that are comparable to the best Nb_3Sn strands and useful critical current densities ($>100\,\text{A}\,\text{mm}^{-2}$ nonstabilizer) can be achieved above 24 T. Perhaps the most important feature of Nb_3Al is the high strain tolerance of the Nb_3Al strands compared to Nb_3Sn. One ton (total length of 210 km) of mass-produced "jelly-roll" process strand (Yamada *et al.* 1994) was manufactured for the 13 T–46 kA Nb_3Al insert for the ITER Central Solenoid model coil. The jelly-roll of Nb and Al is created by co-winding sheets of Nb and Al together on an oxygen-free pure copper core. This will eventually become an individual filament in the final multifilamentary strand. The assembled jelly-roll and core are inserted into an oxygen-free pure copper tube to create an extrusion billet. After hydrostatic extrusion, the billet is drawn down to a small enough size that it can be restacked to create a multifilamentary billet. The billet is hydrostatically extruded and drawn for a second time, to the desired strand diameter size. The extrusion and restack process reduces the thickness of the Nb–Al layers in the filaments to less than 1 μm. The production yield rate for the ITER insert strand was 70% for strands longer than 1.5 km. The final strand is heat treated (800 °C for 10 h is typical) to produce Nb_3Al. The critical current densities obtained by this technique are $\sim 600\,\text{A}\,\text{mm}^{-2}$ at 12 T and 4.2 K (Hosono *et al.* 2002).

The highest critical fields and critical temperatures for Nb_3Al are based around high-temperature heat treatments and rapid cooling. In the rapid-heating, quenching, and transformation (RHQT) (Iijima *et al.* 1997) process Nb/Al composites (MJR or RIT) are resistance heated to 1900–2000 °C and then rapidly quenched into a molten Ga bath at 40–50 °C producing a metastable b.c.c. supersaturated solid solution, $Nb(Al)_{ss}$, inside an Nb matrix. The heating time is ~ 0.1 s. The $Nb(Al)_{ss}$ has excellent ductility and the strand can consequently be mechanically clad or cabled with Cu at this stage and the stabilized strand may also be wound into device form. The $Nb(Al)_{ss}$ is then transformed to Nb_3Al at 800 °C. Both the critical temperature, T_c, of ~ 17.8 K and the critical magnetic field, B_{c2}, of ~ 26 T are 0.7 K and 4 T lower for RHQT composites than the highest values obtained by laser or electron beam irradiation processed

Nb_3Al (Takeuchi *et al.* 2000) and this difference has been attributed (Kikuchi *et al.* 2001a) to stacking faults formed in the A15 phase. Two successful modifications to the basic RHQT technique specifically address the stacking fault issue. In the double rapidly heating/quenching technique (DRHQ), the standard RHQT process is followed, but instead of a conventional low temperature ($<1000\,^{\circ}\mathrm{C}$) transformation heat treatment, a second RHQ process produces the A15 phase and this is followed by a long-range ordering heat treatment at 800 °C for ~ 12 h (Kikuchi *et al.* 2001b). The T_c for these stacking fault free strands has been measured at 18.4 K, which suggests very good stoichiometry. Furthermore, a critical current density of $135\,\mathrm{A\,mm^{-2}}$ was obtained at 25 T (Kikuchi *et al.* 2002). A disadvantage of this technique is that the brittle A15 phase is formed during high-temperature processing. Alternatively, the transformation-heat-based up-quenching technique (TRUQ), follows the initial RHQ of the RHQT process with a short heating step to 1000 °C. At this temperature, the heat released by local transformation from the $Nb(Al)_{ss}$ to the A15 phase produces a rapid local heating and propagation of the reaction along the length of the strand. As the self-heating is localized to the A15 volume, the heat is quickly dissipated after reaction across the Nb and stabilizer volumes, so that external Cu stabilizer applied after the initial RHQ process is not melted (Takeuchi *et al.* 2000).

3.4 HTS Superconductors

HTS superconductors have demanded great interest because of their ability to superconduct with relatively inexpensive liquid nitrogen cooling and their impressive high-field performance. HTS wires and tapes have not achieved the same commercial impact as LTS strands, however, because of the relative difficulty of producing long, low-cost lengths of conductor. The primary sources of this difficulty are the anisotropy in their superconducting behavior and the barrier to current flow presented by high-angle grain boundaries in these materials. Consequently, the strands must be fabricated with a remarkable degree of grain alignment. Further complicating the process is a high sensitivity to heat treatment temperature ($<1\,^{\circ}\mathrm{C}$ control must be exerted across the entire strand length for Bi-2212) and a need for oxygen diffusion into the HTS composite during fabrication. This second requirement effectively means that silver must be used as the superconducting strand stabilizer. Of the HTS superconductors, only Bi-2212 can be readily made into round-wire form strand. All the other HTS superconductors require a tape format. $(Bi,Pb)_2Sr_2Ca_2Cu_3O_x$ based conductors have been shown to be capable of application to superconducting power cables, magnetic energy-storage devices, transformers, fault current limiters and motors, but widespread application will require significant cost reduction in processing. YBCO-based superconductors have the highest recorded critical current densities, but at present there is no viable wire production technique available for YBCO and critical current densities fall rapidly with superconductor thickness above 0.5 μm and with length beyond 1 μm.

(*a*) *Bi-2223.* $(Bi,Pb)_2Sr_2Ca_2Cu_3O_x$ ("Bi-2223") is most readily fabricated using PIT techniques very similar to those used for Nb_3Sn, with Ag replacing the Cu tubes used for Nb_3Sn. In addition to standard wire drawing techniques, groove rolling is also incorporated to bring powder-filled monofilaments to restack size. Hot extrusion can then be used to bond the multifilamentary composite while increasing the powder density and reducing the billet diameter. Extrusion is followed by further drawing and groove rolling, and finally rolling into tape. As a conventionally mechanically processed material, this is readily scalable to large production quantities. The key issue is to increase the connectivity between the grains of the conductor. The prospect of two- to fivefold improvements in the performance of the present process, coupled to cost reductions triggered by significant scale-up and alternative routes such as overpressure processing (Rikel *et al.* 2001) or melt processing (Flükiger *et al.* 2001) could accelerate the early penetration of superconducting technology into the utility network.

Demonstration HTS transmission cable projects have utilized Bi-2223 strand including test installations at Detroit Edison and Southwire in the USA. For the Detroit–Edison project, a conductor was fabricated by American Superconductor that used four layers of Ag-sheathed Bi-2223 tapes (29 km in total), each tape having an average critical current of ~ 118 A d.c. at 77 K. The tapes utilized oxide-dispersion-strengthened silver matrices as well as stainless steel reinforcement added to both sides of the HTS tape (Masur *et al.* 2001). For the Southwire project, 30 m cables were produced from PIT Bi-2223 tapes manufactured by IGC. The measured critical current for the cables was 2980 A at 77 K (Sinha *et al.* 2001).

(*b*) *MgB_2.* Since its discovery in early 2001, MgB_2 has generated much interest as it has the highest T_c (39 K) of any intermetallic superconductor and has the lowest cost components. Like Bi-2223, MgB_2 wires or tapes can use essentially the same technology as that used for Nb_3Sn PIT composites. PIT offers a scaleable technology for future production of kilometer-length MgB_2 and has already been exploited for strand fabrication (Jin *et al.* 2001, Grasso *et al.* 2001). The perpendicular irreversibility field is less than that for Nb_3Sn, the most widely used intermetallic superconductor, and this may restrict MgB_2 to applications of lower field.

4. Summary

Nb–Ti and Nb_3Sn superconductors are well established as commercial superconductors with stable markets for NMR and MRI devices as well as high magnetic field testing equipment. Future markets in fusion reactors, accelerator magnets, and energy storage drive further developments. The strain tolerance of Nb_3Al should maintain interest in this material for high-field applications where conductor cost is not an issue. The ability to use liquid nitrogen as a coolant for HTS conductors, such as Bi-2223, is important for applications such as power transmission lines that are not well suited for low temperature refrigeration. Further performance improvement and cost reduction will be required before HTS strands are economically viable. MgB_2 offers the potential of a low-cost superconductor for the future that, while not being able to use liquid nitrogen cooling, is still able to benefit from reduced refrigeration costs due to its high T_c.

Bibliography

Cave J R, Fevrier A, Verhaege T, Lacaze A, Laumond Y 1989 Reduction of AC loss in ultra-fine multifilamentary NbTi wires. *IEEE Trans. Magn.* **25**, 1945–8

Cheng-ren Li, Xiao-zu Wu, Nong Zhou 1983 NbTi superconducting composite with high critical current density. *IEEE Trans. Magn.* **19**, 284–7

Flükiger R, Giannini E, Lomello-Tafin M, Dhallé M, Walker E 2001 Phase formation in Bi,Pb(2223) tapes. *IEEE Trans. Appl. Superconduct.* **11**, 3393–8

Ghosh A K, Sampson W B, Gregory E, Kreilick T S 1987 Anomalous low field magnetization in fine filament NbTi conductors. *IEEE Trans. Magn.* **23**, 1724–7

Grasso G, Malagoli A, Ferdeghini C, Roncallo S, Braccini V, Siri A S, Cimberle M R 2001 Large transport critical currents in unsintered MgB_2 superconducting tapes. *Appl. Phys. Lett.* **79**, 230–2

Gregory E, Kreilick T S, Wong J, Ghosh A K, Sampson W B 1987 Importance of spacing in the development of high current densities in multifilamentary superconductors. *Cryogenics* **27**, 178–82

Hlasnik I, Takacs S, Burjak V P, Majoros M, Krajcik J, Krempasty L, Polak M, Jergei M, Korneeva T A, Mironova O N, Ivan I 1985 Properties of superconducting NbTi superfine filament composites with diameter $\leq 0.1\,\mu m$. *Cryogenics* **25**, 558–64

Hosono F, Iwaki G, Kikuchi K, Ishida S, Ando T, Kizu K, Miura Y, Sakasai A 2002 Production of a 11 km long jelly roll processed Nb_3Al strand with high copper ratio of 4 for fusion magnets. *IEEE Trans. Appl. Superconduct.* **12**, 1037–40

Iijima Y, Kosuge Y M, Takeuchi T, Inoue K 1997 Nb_3Al multifilamentary wires continuously fabricated by rapid-quenching. *Adv. Cryogenic Eng.* **42**, 1447–54

Jin S, Mavoori H, Bower C, van Dover R B 2001 High critical current in iron-clad superconducting MgB_2 wires. *Nature* **411**, 563–5

Kanithi H, Erdmann M, Valaris P, Krahula F, Lusk R, Gregory E, Zeitlin B 1990 A status report on the development of inner and outer conductors for the SSC dipole and quadrupole magnets. In: McAshan M (ed.) *Supercollider 2*. Plenum, New York, pp. 601–9

Kato T, *et al.* 2001 First test results for the ITER central solenoid model coil. *Fusion Eng. Design* **56–57**, 59–70

Kikuchi A, Iijima Y, Inoue K 2001a Microstructures of rapidly-heated/quenched and transformed Nb_3Al multifilamentary superconducting wires. *IEEE Trans. Appl. Superconduct.* **11**, 3615–8

Kikuchi A, Iijima Y, Inoue K 2001b Nb_3Al conductor fabricated by DRHQ (double rapidly-heating/quenching) process. *IEEE Trans. Appl. Superconduct.* **11**, 3968–71

Kikuchi A, Iijima Y, Inoue K, Kosuge M 2002 New Nb_3Al conductor made by DRHQ process. *Adv. Cryogenic Eng.* **48**, 1025–33

Kiyoshi T, Sato A, Takeuchi T, Itoh K, Matsumoto S, Ozaki O, Wada H, Yoshikawa M, Kamikado T, Ito S, Miki T, Hase T, Hamada M, Hayashi S, Kawate Y, Hirose R 2001 Development and operation of superconducting NMR magnet beyond 900 MHz. *IEEE Trans. Appl. Superconduct.* **11**, 2347–50

Kreilick T S, Gregory E, Scanlan R M, Ghosh A K, Sampson W B, Collings E W 1988 Reduction of coupling in fine filamentary Cu/NbTi composites by the addition of manganese to the matrix. *Adv. Cryogenic Eng.* **34**, 895–900

Lee P J, Larbalestier D C 1993 An examination of the properties of SSC Phase II R and D strands. *IEEE Trans. Appl. Superconduct.* **3**, 833–41

Lee P J, McKinnell J C, Larbalestier D C 1989 Progress in the understanding and manipulation in high J_c Nb–Ti alloy composites. In: Murakami Y (ed.) *Proc. New Developments in Applied Superconductivity*. World Scientific Press, pp. 357–62

Lee P J, McKinnell J C, Larbalestier D C 1990 Restricted novel heat treatments for obtaining high J_c in Nb–46.5 wt.%Ti. *Adv. Cryogenic Eng. (Mater.)* **36**, 287–94

Lefranc G, Muller A 1976 Effect of copper additions to superconducting niobium–tin sinter material. *J. Less Common Metals* **45**, 339–42

Lindenhovius J L H, Hornsveld E M, den-Ouden A, Wessel W A J, ten-Kate H H J 2000 Powder-in-tube (PIT) Nb_3Sn conductors for high-field magnets. *IEEE Trans. Appl. Superconduct.* **10**, 975–8

Masur L, Parker D, Tanner M, Podtburg E, Buczek D, Scudiere J, Caracino P, Spreafico S, Corsaro P, Nassi M 2001 Long length manufacturing of high performance BSCCO–2223 tape for the Detroit Edison Power Cable Project. *IEEE Trans. Appl. Superconduct.* **11**, 3256–60

Meingast C, Lee P J, Larbalestier D C 1989 Quantitative description of a high J_c Nb–Ti superconductor during its final optimization strain: I. Microstructure, T_c, H_{c2} and resistivity. *J. Appl. Phys.* **66**, 5962–70

Ochiai S, Uehara T, Osamura K 1986 Tensile strength and flux pinning force of superconducting Nb_3Sn compounds as a function of grain size. *J. Mat. Sci.* **21**, 1020–6

Patnaik S, Cooley L D, Gurevich A, Polyanskii A A, Jiang J, Cai X Y, Squitieri A A, Naus M T, Lee M K, Choi J H, Belenky L, Bu S D, Letteri J, Song X, Schlom D G, Babcock S E, Eom C B, Hellstrom E E, Larbalestier D C 2001 Electronic anisotropy, magnetic field-temperature phase diagram and their dependence on resistivity in *c*-axis oriented MgB_2 thin films. *Supercond. Sci. Technol.* **14**, 315–9

Poole C P, Farach H A 2000 Tabulations and correlations of transition temperatures of classical superconductors. *J. Superconduct.* **13**, 47–60

Rikel M O, Williams R K, Cai X Y, Polyanskii A A, Jiang J, Wesolowski D, Hellstrom E E, Larbalestier D C, DeMoranville K, Riley G N Jr. 2001 Overpressure processing Bi2223/Ag tapes. *IEEE Trans. Appl. Superconduct.* **11**, 3026–9

Sakamoto H, Higuchi M, Endoh S, Kimura A, Wada K, Meguro S, Ikeda M 2000 Very high critical current density of bronze-processed $(Nb,Ti)_3Sn$ superconducting wire. *IEEE Trans. Appl. Superconduct.* **10**, 971–4

Scanlan R M, Fietz W A, Koch E F 1975 Flux pinning centres in superconducting Nb_3Sn. *J. Appl. Phys.* **46**, 2244–9

Scanlan R, McInturff A D, Taylor C E, Caspi S, Dell'Orco D, Higley H, Gourlay S, Bossert R, Brandt J, Zlobin A V 1997 Design and fabrication of a high aspect ratio cable for a high gradient quadrupole magnet. *IEEE Trans. Appl. Superconduct.* **7**, 936–8

Sinha U K, Lindsay D T, Hughey R L Jr., Stovall J P, Gouge M J, Lue J W, Haldar P, Selvamanickam V, Vo N 2001 Development and test of world's first industrial high temperature superconducting (HTS) power cable. *Power Eng. Soc. Winter Meeting, IEEE* **2**, 442–7

Smathers D B, Leonard D A, Kanithi H C, Hong S, Warnes W H, Lee P J 1996 Improved niobium 47 wt.% titanium composition by iron addition. *Mater. Trans.* (Japanese Institute of Metals) **37**, 519–26

Stekly Z J J, Zar J L 1965 Stable superconducting coils. *IEEE Trans. Nucl. Sci.* **12**, 367–72

Takeuchi T, Banno N, Fukuzaki T, Wada H 2000 Large improvement in high-field critical current densities of Nb_3Al conductors by the transformation-heat-based up-quenching method. *Supercond. Sci. Technol.* **13**, L11–14

Warnes W H, Larbalestier D C 1986 Critical current distributions in superconducting composites. *Cryogenics* **26**, 643–53

Wilson M N 1972 Filamentary composite superconductors for pulsed magnets. In: *Proc. 1972 Appl. Superconduct. Conference, IEEE*, pp. 385–8

Yamada Y, Ayai N, Takahishi K, Sata K, Sugimoto M, Ando T, Takahashi Y, Nishi M 1994 Development of Nb_3Al/Cu Multifilamentary Superconductors. *Adv. Cryogen. Eng.* **40**, 907–14

P. J. Lee
The University of Wisconsin, Madison, WI, USA

Supported Catalysts

Heterogeneous catalysts are widely exploited in many industries, particularly the chemical, petrochemical, agrochemical, and pharmaceutical sectors. However, for their successful use, heterogeneous catalysts have to withstand the conditions within the chemical reactor without attrition or degradation. Although some heterogeneous catalysts can be used in bulk or nonsupported form, many involve materials that cannot be used directly without the aid of an additional material, termed a catalyst support. There are many reasons why a support is required and these include stability of the catalyst phase, e.g., stabilization of small metal particles, as well as cost, i.e., dilution of an expensive catalyst component. In this article the choice of the support material, the types of supported catalysts available, and the characterization of the support–catalyst interaction will be described.

1. Choice of Support

The function of the support is to provide a sound base on which to disperse the active catalyst component. Key features of supports that govern their selection are the required porosity/surface area and the chemical properties, acidity, and basicity. The first factor that has to be taken into account in the selection of a support is the nature of the chemical reaction being catalyzed since this can be used to determine the porosity of the support material. For example, if the thermodynamically preferred product is required (as in the water gas shift reaction (Twigg 1989)) then high-surface-area supports can be considered. If the products of kinetic control are required then low-surface-area supports are generally used, e.g., the partial oxidation of ethene to the epoxide using silver supported on α-Al_2O_3(Bowker 1998). The selection of a specific support material will then depend on the chemical properties required for the support surface which will be specific to any particular reaction. The particular feature that is of crucial importance is the relative concentrations of acidic and basic sites. Support surfaces, particularly oxides, can contain Brønsted and Lewis acidic or basic sites; e.g., surface hydroxyl groups on the surface of γ-Al_2O_3 can act as a Brønsted acid. Typically oxides are used as support materials, with the most common being alumina (α and γ forms), silica, magnesia, titania, zirconia, and zeolites.

2. Types of Supported Catalysts

This section sets out the different types of supported catalysts in current use.

2.1 Zeolites as Supports

Zeolites make excellent candidates for catalyst supports due to their stability, high surface area, and probably most importantly, the regularity of their pore structure (van Bekkum *et al.* 1991, Thomas and Thomas 1997) (see *Zeolites**). The well-defined pore sizes of zeolites means that when used as catalysts or supports they can impose selectivity on the reaction products by restricting access of reactants to the active site, the formation of certain transition states, or the egress of certain products. Thus it is vital when preparing catalysts supported on zeolites that as little active material as possible ends up on the external surface of the zeolite where this selectivity will not be available.

Zeolitic frameworks are negatively charged, and this negative charge is balanced by the presence of extra-framework cations such as Na^+ or H^+. The cationic sites in zeolites can be substituted with metal cations, thus introducing potentially new catalytic species into the structure. However, although metals supported in this form are used as catalysts it is more common for further treatment to be carried out to form the active catalyst. In the majority of cases calcination and reduction are carried out on the metal cations to form neutral metal atoms or groups of atoms which act as the catalytically active site. Zeolites can therefore be used as supports for metal cations that are active catalysts (Thomas and Thomas 1997) (see *Acid Catalysts*[#]).

The most important ways for introducing metal cations into zeolites are ion exchange and the incipient wetness technique. In ion exchange the cations present in the zeolite are substituted by the required metal cation by treating the zeolite with a solution of a metal salt. This can be reduced to form small metal particles if required. The exact nature of all of treatment steps is crucial to the structure of the final product (van Bekkum *et al.* 1991) due to the difference they can make to the final location of the metal, as zeolites generally have several different sizes of pores within their structure. It is important that the introduced metal ends up in the larger pores of the zeolite structure, otherwise access of reactants to the metal, vital for catalytic reactivity, is reduced or even made impossible. There is often a tendency for the metal ions to migrate to the smaller pores due to the highly coordinating nature of the oxygen atoms which can envelop them in such confined areas. This can be avoided by blocking the smaller cages with inactive cations before introduction of the catalytically active metal.

One advantage that zeolites have over other supports is the ability to isolate metal atoms in the pores and so prevent sintering of the metal atoms, which would greatly reduce the effective surface area of the catalyst. However, to prevent agglomeration of the metal it is important to control all the stages in the preparation to avoid atom migration. Other strategies exist for introducing the metal to the zeolite, some of which introduce the metal as part of a neutral complex using either condensation from the vapor phase (CVD) or adsorption from an organic solvent. In this, complexes can be inserted into the zeolite pores and used as catalysts either as they are or with the ligands removed and the metal reduced under milder conditions than would be possible using more conventional methods. It is also possible to introduce more than one metal to the zeolite using any of these methods.

Structured mesoporous amorphous silicas and silica aluminas (denoted MCMs) have been fabricated that contain ordered pore structures with channel apertures of 2.5–10 nm. These can also be used as supports for cations, complexes and metals (Thomas and Thomas 1997).

2.2 Well-defined Oxides for Supporting Active Catalysts Molten at Reaction Temperature

Support of catalysts which are molten at the reaction temperature can be considered as a special case of supported liquid-phase catalysis. The molten catalyst fills all the pores on the support and so, if properly prepared with a support-to-catalyst ratio calculated to give a thin layer of catalyst, will give a high catalyst surface area. The catalytic process itself is in effect homogeneous, with the reactant molecules dissolving in the melt from the gas or vapor phase, reacting in the liquid phase, and then the products diffusing out. Thus the advantages of heterogeneous catalysis (large reaction area, easier reactor design, and ease of separation of products) and those of homogeneous catalysis (high catalyst dispersion and increased specificity) are combined. Advantages over supported liquid-phase catalysis are that catalyst loss due to evaporation at operating temperatures is minimal and also that operating temperatures can be higher.

To obtain a catalyst that will melt at reasonable temperatures (not high enough to destroy thermally the reactants or be too expensive to maintain) there are a number of approaches. If possible a single catalyst component with a suitably low melting point can be used. However, the stable metals and metal salts and oxides used for catalysis rarely have this property. An alternative strategy is to use a eutectic mix, where the combination of two or more materials lowers the melting point of the resulting mix. An example of this is the mixture of Cs_2O, MoO_3, and V_2O_5 (pure oxide melting points 490 °C, 795 °C, and 690 °C, respectively) which melts at 350 °C, described by van Setten *et al.* (1999).

Another more general approach uses a low-melting-point salt as a solvent for the active catalyst components. Chloride salts are often used both to form eutectics and as low-melting ionic solvents (Lago *et al.* 1995). Tetrabutylammonium chloride is an example of a commonly used organic solvent (melting point 44 °C). Obviously the use of such organic salts limits the upper temperature at which the system may be used, as decomposition will occur at elevated temperatures. Thus while inorganic salts such as CuCl/KCl are used at temperatures exceeding 400–500 °C, mixtures containing organic salts are generally not suitable for use above 200 °C.

The choice of support is important as not all supports are suitable for all catalyst mixes. It is important that the molten material "wets" the support thoroughly. For example, while silica has been found to be a good support for many molten catalysts it is unsuitable for the mixture of Cs_2O, MoO_3, and V_2O_5 cited previously.

Characterization of supported molten catalysts can be problematical. Most of the characterization reported in the literature has been done on cooled or quenched samples rather than *in situ* on molten samples. X-ray diffraction data suggest that in the quenched state these systems remain amorphous.

2.3 Supported Oxide Monolayer Catalysts

The use of supported metal oxide catalysts, where the active oxide material is present in monolayer form as illustrated in Fig. 1, can produce significantly different catalyst performance from that obtained for the corresponding bulk metal oxide material (Roozeboom *et al.* 1979). This is also observed when monolayer catalysts are compared to supported oxide catalysts where the active oxide loadings and general preparation methods result in the formation of bulk-like oxide species at the catalyst surface. For example, CuO supported on γ-alumina at monolayer CuO coverage exhibits higher activities for nitrous oxide decomposition, a reaction of considerable environmental importance, in comparison to catalysts prepared by impregnation or mechanical mixing of equivalent CuO loadings (Yao *et al.* 1998). Also, molybdenum oxide at monolayer coverage on alumina displays the highest activity for dehydrogenation of *n*-butane (Harlin *et al.* 1999). Other reactions in which supported metal oxide monolayer catalysts demonstrate superior catalyst performance are selective oxidations, *i*-propanol decomposition, and ammoxidation of aromatics (Bond and Tahir 1991). Most research on monolayer oxide catalysts has involved vanadium oxide species supported on materials such as titania, alumina, zirconia, and ceria. Supported vanadium oxide monolayers have therefore been adapted as the model systems for fundamental research in this area of catalysis.

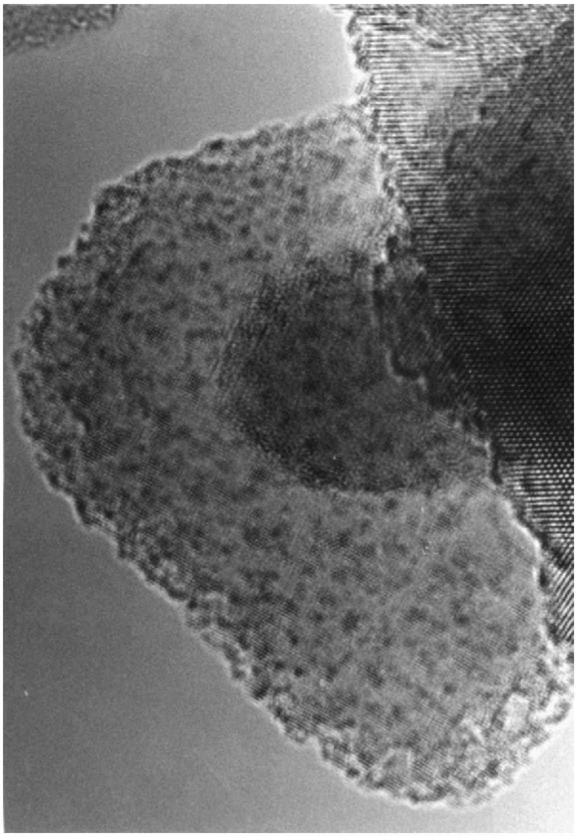


Figure 1
TEM characterization of a WO_3/TiO_2 monolayer catalyst, WO_3 visible on TiO_2 by virtue of mass-contrast (reproduced by permission of Baltzer Science from *Catal. Lett.*, 1996, **39**, 219–31).

The introduction of a support increases the thermal and mechanical stability of the metal oxide or oxo species along with increasing the surface area of the active material. Using a support may also prove to be more economically favorable as less active material is required. The nature of the oxide support can itself be an important factor in determining the catalyst performance and this is demonstrated in the oxidation of *n*-butane to maleic anhydride over vanadium oxide monolayer catalysts (Bond and Tahir 1991). This is related to the extent or strength of interaction between the support and the monolayer species, the bonding within the monolayer itself, and the acid–base properties of the system.

A number of suggestions are given by Yao *et al.* (1998) as to the origin of the differences in catalytic behavior of supported monolayer oxides. It is proposed that changes in the environment around the metal cations due to variations in lattice energy, chemical bonding between the support and active material, or the formation of an epitactic layer can exert significant influence on the catalytic activity.

Several methods have been employed in the preparation of monolayer catalyst systems. The catalyst precursor materials can be formed via impregnation, grafting (adsorption from solution), co-precipitation, or chemical vapor deposition methods. The precursor materials are then calcined at the appropriate temperature to form the active catalyst. Factors such as the preparation conditions (temperature, solvent, and precursor compound), pretreatment of the support material, and calcination conditions can play a major role in determining the uptake of the precursor species and characteristics of the final oxide monolayer.

A number of techniques can be used to characterize monolayer formation and the surface chemistry of these materials relating to structure and monolayer–support interactions. Laser Raman spectroscopy can be used to establish the point at which monolayer coverage is reached, as well as providing important structural information by discriminating between species present at submonolayer coverages and those present at above monolayer coverage, i.e., crystalline oxide material. In the case of supported vanadium monolayers the point of monolayer formation is thought to occur immediately prior to the appearance of a crystalline V_2O_5 phase. Species which indicate lower than monolayer coverage, such as isolated $VO_4{}^{3-}$ tetrahedra and distorted vanadate octahedra, can be identified by the presence of bands at 830 and 970 cm^{-1} respectively (Bond and Tahir 1991).

2.4 Supported Metal Catalysts

Supported metal catalysts are used extensively in the chemical industry, from bulk and fine chemicals production to petrochemicals. A wide range of reactions can be performed effectively over supported metal catalysts such as hydrogenation, hydrogenolysis, dehydrogenation, reforming reactions, and cracking of petroleum hydrocarbons. The main active components in these systems are typically Group 10 or 11 metals.

Although a number of industrial processes employs unsupported metal catalysts, e.g., Raney nickel in the hydrogenation of glucose to sorbitol and platinum–rhodium gauze in the oxidation of ammonia in the manufacture of nitric acid, the use of a support has many advantages. Firstly, smaller amounts of the often expensive active metal component are required as the metal is highly dispersed on the support material as small metal particles. The use of a support also increases the available metal surface area which can be directly related to the catalytic activity. The support therefore provides a structural framework and stabilizes the active species (small metal particles) as well as making the catalyst more mechanically robust. In some cases the support itself can introduce new active sites to the system, which are usually acidic or basic or both, and a polyfunctional catalyst is obtained.

Typical support materials include high-surface-area, porous, metal oxides (e.g., Al_2O_3, SiO_2, and TiO_2), zeolites, and carbons. In some cases the support only plays a physical role and is considered inert towards the catalytic process. However, depending on the material used, the support can influence the electronic environment of the active sites, change the number of active sites, and exert influence on the morphology of the metal crystallites. When this occurs, metal–support interactions are in operation and the strength of these interactions can have significant effects on the catalytic performance (Bond 1982, Augustine 1995).

In general, when reducible metal oxides are employed as the support the metal–support interactions are strong, particularly after high-temperature reduction pretreatment (Bond 1982). For example, with TiO_2, this reduction treatment causes reduced TiO_x species to migrate across the metal surface and this results in the creation of new active sites associated with Ti cations or oxygen vacancies. This phenomenon is known as a strong metal–support interaction, termed SMSI. This is not observed with nonreducible metal oxide carriers such as SiO_2 and zeolites although relatively weak metal–support interactions can be prevalent which can still have considerable or at least observable effects on the catalyst behavior.

In most industrial applications small amounts of additives or "promoters" are present in the supported metal catalyst system and these can have a number of functions which improve the catalyst efficiency (Twigg 1989). For example, the introduction of promoters can increase both the number and intrinsic activity of active sites.

Most supported metal catalysts are susceptible to poisoning by molecules containing sulfur, nitrogen, chlorine, and oxygen, even at low concentrations, and this can have serious effects on large-scale industrial applications. However, in certain cases (Hutchings *et al.* 1994) controlled doping of supported metal catalysts with small amounts of such molecules can improve the catalyst selectivity and therefore act as promoters. The action of promoters and poisons can usually be described in terms of electronic or geometric effects.

The main techniques utilized in the preparation of supported metal catalysts are impregnation and co-precipitation methods. A number of other techniques may also be used including ion exchange (particularly for zeolite-supported metals), deposition (precipitation–deposition and reduction–deposition), and metal vapor deposition (Bond 1982, Che and Bennet 1989).

Impregnation involves contacting the support material with a solution of suitable metal salt, e.g., nitrate, under controlled conditions over a period of time to allow adsorption of the catalyst precursor salt from the solution. The volume of solution can either be equal to the total pore volume of the support (incipient wetness) or in excess of the pore volume. Incipient wetness is generally used when reproducible metal loadings are required and when there are only weak interactions between the support and the precursor material. A number of variables is important in impregnation procedures such as the concentration of the precursor salt solution, the time of contact, the solvent, and the temperature.

In the case of co-precipitation, the metal precursor and support materials are formed simultaneously by addition of a basic precipitating agent (usually carbonate or hydroxide) to a solution containing appropriate metal and support precursor salts. This results in the distribution of the metal component throughout the bulk structure of the material, whereas with impregnation the active metal species are deposited on the surface only. The latter may be more economically favorable since a larger quantity of metal salt will be required to give an equivalent surface metal loading, and therefore active surface area, than with the co-precipitated method.

After the catalyst precursors have been formed and dried the samples are then usually calcined in air to decompose the precursor species to the oxide as these can often be more easily reduced to the metallic state. Calcination is followed by reduction in a gas such as hydrogen or carbon monoxide under controlled conditions. In some cases, it is more favorable to omit the calcination step and reduce the catalyst precursor directly. Sometimes during calcination highly stable compounds are formed which are difficult to reduce under typical reduction conditions, e.g., metal aluminates and silicates.

The choices of preparation method and pretreatment conditions are important in determining the performance of the catalyst. These factors can have direct effects on metal surface area, dispersion (ratio of number of metal atoms exposed at the surface to total number of metal atoms present), and in particular the metal particle size and morphology. These are important parameters which are directly related to catalyst activity, selectivity, and stability and therefore need to be characterized in order to obtain accurate and quantitative measurement of catalyst performance.

3. Characterization of Support–Catalyst Interaction

A vast array of techniques can be used to characterize supported catalysts with the aim of gaining information concerning the metal surface area, dispersion and particle size, oxidation states of surface species, metal–support interactions, and the factors which influence these characteristics. In general, measurement of active surface area and dispersion is achieved using selective chemisorption. Adsorption of a gas (e.g., hydrogen, oxygen, or carbon monoxide) on the metal particles alone can be used to calculate the number of metal atoms exposed at the surface of the catalyst, providing the adsorption stoichiometry (number of gas molecules or atoms adsorbed per metal atom) is known. Metal particle size can also be obtained by means of electron microscopy, e.g., transmission electron microscopy (TEM), x-ray line broadening using the Scherrer equation, as well as magnetic methods. Electron microscopy is particularly useful as the shapes of the metal particles can be observed visually and values for dispersion and particle size distributions obtained. *In situ* Fourier transform infrared (FTIR) spectroscopy can be used to determine the oxidation states of the metal species present at the surface by means of CO adsorption studies. X-ray photoelectron spectroscopy (XPS) can also be used to establish oxidation states, and as with FTIR spectroscopy vital information on metal–support interactions can be obtained, such as the nature and strength of the bonding between the metal component and the support. Other techniques which can be used to study the nature of metal–support interactions include NMR (nuclear magnetic resonance) spectroscopy, UPS (ultraviolet photoelectron spectroscopy), Mössbauer spectroscopy, and EXAFS (extended x-ray absorption fine structure) as well as monitoring the effects of the support on chemisorptive and catalytic properties. To date, the best methods for studying the nature of the support–catalyst interactions are TEM, HREM (high-resolution electron microscopy) (Thomas and Thomas 1997), and STM (Bowker 1998).

Bibliography

Augustine R L 1995 *Heterogeneous Catalysis for the Synthetic Chemist*. Marcel Dekker, New York

Bond G C 1982 *Metal-support and Metal-additive Effects in Catalysis*. Elsevier, Amsterdam

Bond G C, Tahir S F 1991 Vanadium oxide monolayer catalysts, characterization and catalytic activity—a review. *Appl. Catal.* **71**, 1–31

Bowker M 1998 *The Basis and Applications of Heterogeneous Catalysis*. Oxford University Press, Oxford

Burrows A, Kiely C J, Joyner R W, Knozinger H K, Lange F 1996 An HREM study of WO_3/TiO_2 monolayer catalyst system. Proposals for the overlayer structure. *Catal. Lett.* **39**, 219–31

Che M, Bennet C O 1989 The influence of particle size on the catalytic properties of supported metals. *Adv. Catal.* **36**, 55

Harlin M E, Backman L B, Krause A I O, Jylhä O J T 1999 Activity of molybdenum oxide catalyst in the dehydrogenation of *n*-butane. *J. Catal.* **183**, 300

Hutchings G J, King F, Okoye I P, Padley M B, Rochester C H 1994 Selectivity enhancement in the hydrogenation of α,β unsaturated aldehydes and ketones using thiophene-modified catalysts. *J. Catal.* **148**, 453

Lago R M, Green M L H, Tsang C S, Odlyha M 1995 Catalytic decomposition of chlorinated organics in air using copper chloride-based catalysts. *Appl. Catal. B* **8**, 107–21

Roozeboom F, Fransen T, Mars P, Gellings P J 1979 Vanadium oxide monolayer catalysts. 1. Preparation, characterization, and thermal stability. *Z. Anorg. Allg. Chem.* **25**, 449

Thomas J M, Thomas W J 1997 *Principles and Practice of Heterogeneous Catalysis*. VCH, Weinheim

Twigg M V 1989 *Catalyst Handbook*. Wolfe, London

van Bekkum H, Flanigen E M, Jansen J C 1991 *Introduction to Zeolite Science and Practice*. Elsevier, Amsterdam

van Setten B A A L, van Dijk R, Jelles S J, Makkee M, Moulijn J A 1999 The potential of supported molten salts in the removal of soot from diesel exhaust gas. *Appl. Catal. B* **21**, 51–61

Wu X, Letuchy Y A, Eyman D P 1996 Catalytic dechlorination of CCl_4 over silica-supported $PdCl_2$-containing molten salt catalysts: the promotional effects of $CoCl_2$ and $CuCl_2$. *J. Catal.* **161**, 164–77

Yao K-W, Jaenicke S, Lin J-Y, Tan K L 1998 Catalytic decomposition of nitrous oxide on grafted CuO/γ-Al_2O_3 catalysts. *Appl. Catal. B* **16**, 291–301

J. E. Bailie, G. J. Hutchings and S. O'Leary
Cardiff University, UK

T

Textile Fibers: A Comparative Overview

Fibers from natural sources, twisted by hand into yarns, and then woven into textile fabrics, constitute a materials technology which dates back over 10 000 years. Apart from hand-tools, the technology changed little until the industrial revolution, with the invention of power machines concentrated in 1775–1825. The available fibers, namely cotton, some fibers extracted from the stems or leaves of plants, wool, other hairs, and silk, remained unchanged for another 100 years. All except silk were short fibers, with staple lengths of about 1–10 cm. These fibers had to be twisted into yarns. Even silk filaments were of finite length and had to be "thrown" together into longer yarns, which were smooth and lustrous in contrast to the hairy staple-fiber yarns. Advances in chemistry led to solutions of cellulose derivatives, which could be extruded through multiple holes, coagulated, and regenerated as continuous filament yarns of effectively infinite length. For a time, these were known as *artificial silk*. The most successful, *viscose rayon*, was first commercially produced in 1905. The recognition of the idea of macromolecules in the 1920s led to manufactured, synthetic yarns of several vinyl polymers, but the major invention was nylon, which became commercial in 1938. Polyester followed 10 years later. The two other major synthetic fibers of this first generation were acrylics and polypropylene. A second generation of high-performance fibers started with the aramid fiber, *Kevlar*, followed by high-modulus polyethylene. Elastomeric fibers, such as *Lycra*, were another development.

The uses of textile fibers fall into three categories. Both *clothing* and *furnishing fabrics* are somewhat unusual in materials technology, since color and the other esthetic features of pattern and feel, which are determined by the fiber and textile structures, are as important as the functional requirements for cover, protection, warmth, and durability. The third category of *technical textiles* includes some old and simple uses, ranging from ropes to wiping cloths, but is becoming of increasing importance in demanding engineering and medical applications.

All the fibers mentioned above are polymeric materials, and they, with some other polymer fibers of more limited or specialist interest, comprise the main scope of this article. Inorganic fibers, namely glass, ceramic, carbon, and metallic, although they can be formed into textile structures, will only be mentioned briefly.

[The sizes of fibers are most conveniently expressed in terms of linear density, namely mass per unit length, rather than area of cross-section or diameter. The SI unit is the tex, which is grams per kilometer (g km^{-1}). The submultiple decitex (dtex) is widely used; an older unit, denier, is grams per 9000 m. This leads to the mechanical quantities specific stress, modulus, and strength (tenacity) being expressed as (force/linear density). The preferred unit is newtons per tex (N tex^{-1}), which is dimensionally equivalent and equal to gigapascals per gram per cubic centimeter (GPa g^{-1} cm^{3}), kilojoules per gram (kJ g^{-1}), or, in a link to wave propagation velocity, the square of kilometers per second (km s^{-1})^{2}.]

1. *Fibers and Fiber Types*

1.1 *Fibers as a Material Form*

Fibers, which are materials in a one-dimensional form characterized by flexibility, fineness, and high ratio of length to thickness (McIntyre and Daniels 1995), perform many functions in living organisms and are used in a wide variety of manufactured structures. A list of values of the fiber form, some of which are more relevant to other fiber uses than to textiles, includes:

(i) Combination of flexibility with strength, in contrast to the usual association of stiffness and strength.

(ii) Crack-stopping at the discontinuities, which gives strength to composites.

(iii) High surface area, which is important in absorption.

(iv) Large included volume between fibers in textile structures.

(v) Continuity over long lengths, which is vital in optical fibers.

(vi) Ability to form networks.

(vii) Structural control: in natural fibers, by genetics; in manufactured fibers, due to the rapid heating, cooling, evaporation, and stress changes.

(viii) Small defect size (less than fiber diameters), typified by the high strength of glass fibers.

(ix) Ability to modify chemistry or introduce additives to give specific properties.

(x) Control of end-use performance at several levels: chemical constitution, fiber fine structure, macroscopic fiber form, yarn structure, fabric structure, etc.

1.2 *Characteristic Features of Textile Fibers*

For clothing and household textiles, the common fiber requirements are:

Appropriate dimensions: commonly, textile fibers are in the range 1–20 dtex, which gives diameters of 5–50 μm. Coarser forms, typically 0.1–1mm, are known as monofilaments. *Microfibers*, down to 0.1dtex ($\sim$3 μm diameter) or less, have been introduced in recent years. *Nanofibers*, with linear densities from 10^{-3}–10^{-7}dtex (ca. 3–300 nm), have been produced by electro-spinning, which involves the extrusion of polymer solutions or melts in a high electric field (Reneker and Chun 1996).

---

* Cross references marked by an asterisk are included in this volume.

# Cross references marked by a hash can be found by consulting the Encyclopedia of Materials: Science and Technology.

Mechanical, thermal, and chemical stability in everyday use.

Accessibility to dyestuffs for coloration.

An intermediate range of elastic extensibility; break extensions from 5 to 50%.

The only materials that have been found to satisfy these requirements for natural or manufactured fibers are *partially oriented*, *partially crystalline*, *linear polymers*. For special purposes, fibers outside this range of requirements may be polymeric or inorganic. High-stretch, elastomeric fibers have break extensions up to 500%, and high-modulus, high-tenacity (HM-HT) fibers down to 2%. Coloration is not needed for engineering fibers. Higher temperature or chemical resistance or other special properties are needed for some applications.

1.3 Chemical and Manufacturing Classification and Importance

The general textile fiber market is satisfied by six chemical types:

(i) cellulose, in cotton, other plant fibers, and rayon and its derivative cellulose acetate;

(ii) proteins in wool, hairs, and silk;

(iii) polyamides, mainly nylon 6 and 66;

(iv) polyesters, mainly polyethylene terephthalate (PET or 2GT);

(v) polyacrylonitrile (PAN) acrylic fibers, and other vinyl polymers;

(vi) polyolefins, mainly polypropylene.

In addition, the elastomeric fibers are polyurethanes. Aramids (aromatic polyamides) and other polymers are used as HM-HT fibers, for high-temperature resistance, and for other special purposes. Table 1 gives statistics, which show the relative importance of the different materials, with some indication of the division between different forms and applications of the same material.

Cotton, other natural cellulose fibers, wool, and hairs are laid down over weeks in living cells. Silk is extruded as an aqueous solution stream, which dries into the filaments. In spider silk, the rate of extrusion may be as high as 10 cm min^{-1}. Manufactured polymer fibers are produced at rates of 100–10 000 m min^{-1}. There are several methods of production. From polymer solutions, either *wet-spinning* into a coagulating bath, sometimes involving a chemical reaction, or *dry-spinning* with evaporation may be used. Thermoplastic polymers enable the faster *melt-spinning* process to be used. High-modulus polyethylene (HMPE) fibers are made by *gel-spinning*.

Table 1
Textile fiber statistics.

Fiber type	Annual world demand or production 1997 (× 1000 t)
Natural	
Cotton	19453[a]
Other plant fibers	4648[b]
Flax	633
Sunn hemp	81
True hemp	27
Jute and jute-like	3266
Sisal and henequen	342
Others	299
Wool	1429[a] (scoured 840)
Silk	86[a]
Manufactured	
Regenerated cellulosic	2309[a] (capacity 3172)
Viscose	
Industrial yarn	80
Regular yarn	318
Staple	1220
Acetate	
Yarn	107
Cigarette tow	584
Synthetic polymer	21683[a]
World capacity	
PET	
Yarn	10810
Staple	9383
Nylon	
Yarn	4843
Staple	706
Acrylic	
Yarn	6
Staple	3286
Others (except olefin)	
Yarn	201
Staple	190
World production	
Olefin	
Yarn	1696
Staple	1012
Slit film	1891
US production	
Polyester	1844
Industrial yarn	202
"Textile" yarn	546
Staple	1097
Nylon	1288
Industrial yarn	130
Carpet yarn	609
"Textile" yarn	188
Staple	362
Glass	2570

[a] a *Demand from Fiber Organon* June 1997. [b] b Production from FAO, *FAO Yearbook*, *Production*, 19??, Food and Agriculture Organization of the United Nations, Rome, Vol. 51. Other figures estimated from other sources.

Inorganic fibers are dealt with in detail in other articles, particularly in relation to composites. Glass fibers have been known since ancient times, and became of major importance, manufactured in various ways, in the twentieth century. They are occasionally used in clothing or furnishing fabrics, but their major uses are in technical textiles and composite preforms (see *Glass Fibers**). Optical fibers may be of glass or organic polymers (see *Glass Optical Fibers*#). Low-grade ceramic fibers, used for insulation, are made by melt processes. High-performance ceramic fibers, such as those based on silicon carbide or alumina, may be formed as polymer fibers, with the inorganic elements either incorporated in the polymer or mixed as compounds in the polymer solution, and then heat treated for conversion to ceramic; their main use is in composites (see *Ceramic Fibers from Polymer Precursors**; *Spun (Slurry and Sol–Gel) Ceramic Fibers**). Other ceramic fibers, such as boron or silicon carbide, may be made by chemical vapor deposition (see *CVD Monofilaments**). The highest strengths are achieved with single-crystal whiskers (see *Whiskers**). Most carbon fibers (see *Carbon Fibers**) are made by heat treatments of acrylic fibers, but rayon can also be used, and pitch-based fibers have somewhat different structure and properties. Moderate heat treatment gives thermally resistant fibers used in protective clothing, and higher temperature treatments give high-strength, high-modulus fibers, which are mainly used in composites (see *Fibrous Reinforcements for Composites: Overview**) but can also be used, for example, in medical textiles. Fine metal wires can be regarded as fibers, and may be used for decorative or physical reasons (see *Metallic Filaments**).

1.4 Comparative Properties

Table 2 lists typical properties of the more important textile fibers. Where two sets of mechanical properties are given, these are typical of different types, but the available range may be greater. Figure 1 shows stress–strain curves. These values are for "dry" fibers, namely fibers in equilibrium at a standard atmosphere of 65% relative humidity and 20 °C. Roughly, the fibers can be divided into five groups: *low strength*, *low extension*, e.g. rock wools; *high strength*, *low extension*, e.g. aramid and HMPE; *tough*, due to combined strength and extension, e.g. nylon, polyester, and polypropylene; *less tough*, lower combined strength and extension, e.g. cotton, wool, and rayon; and *very high extension*, e.g. *Lycra*. However, for the natural fibers, properties vary with variety and growth conditions, and, for manufactured fibers of any given type, the degree of orientation imposed allows for a range from lower strength/higher extension to higher strength/lower extension.

The densities of textile fibers are typically in the range 1–1.5 g cm^{-3}, which gives them a considerable weight advantage compared to many structural materials. Cellulose, protein, and, to a lesser extent, nylon absorb appreciable quantities of water, which causes swelling and changes properties. When dry, the polymeric textile fibers are insulators. As water is absorbed the dielectric constant increases, reaching high values at high humidities; this allows ions to dissociate, and lowers the specific resistance to about $10^3\,\Omega$ cm in moisture-absorbing fibers at high humidity. The addition of additives to fibers gives special

Table 2
Typical properties of textile fibers.

Fiber	Density (g cm^{-3})	Melting point (°C)	Moisture at 65% relative humidity, 20 °C (%)	Tenacity (N tex^{-1})	Break extension (%)	Work of rupture (J g^{-1})	Initial modulus (N tex^{-1})
Cotton	1.52	~200(chars)	7–8	0.19	5.6	5.1	3.9
				0.45	6.8	15	7.3
Viscose rayon	1.49	~200(chars)	12–14	0.18	27	31	4.8
				0.41	12	28	8.8
Acetate (secondary)	1.32	250	6–7	0.13	24	22	3.6
Wool	1.30	130(decomposes)	14–18	0.14	43	38	2.1
Silk	1.34	175(chars)	10	0.38	23	60	7.3
Nylon 6/66	1.14	215/260	4	0.29	46	77	0.6
				0.84	20	100	6–9
PET (polyester)	1.39	260	0.4	0.47	37	119	8.8
				0.82	13	60	10–17
Acrylic	1.19	200(sticks)	1–2	0.27	25	47	6.2
Polypropylene	0.91	165	0	0.65	17	71	7.1
Para-aramid	1.44	500(decomposes)	7–1.2	2.1–1.6	4.4–2.5	45–30	51–98
HMPE	0.97	147	0	2.6–3.5	3.8–2.7		90–170

From Morton and Hearle (1993), and manufacturers' data.

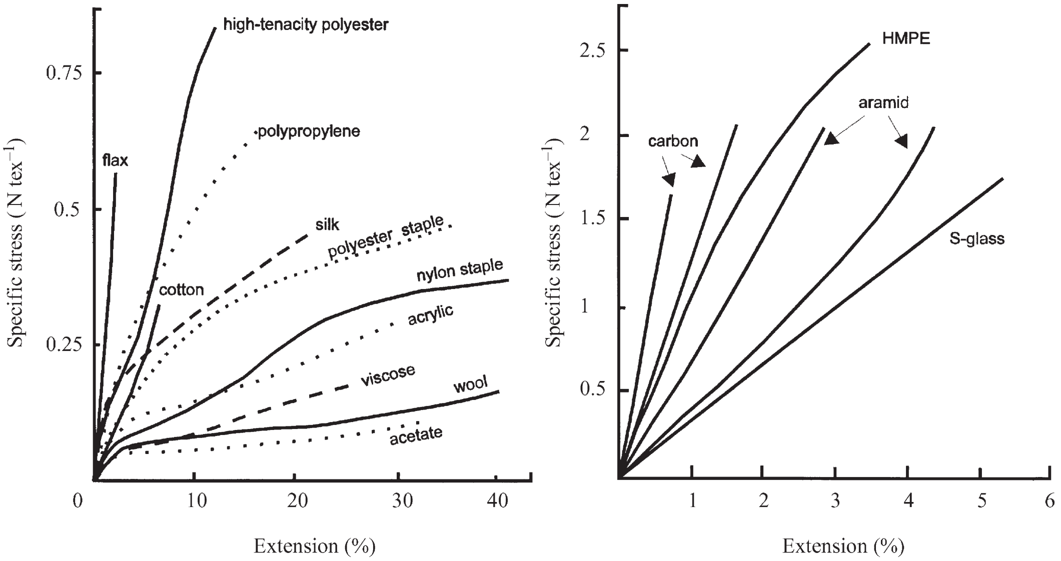


Figure 1
Typical stress–strain properties of textile fibers at 65% relative humidity and 20°C. (a) Natural fibers and first-generation manufactured fibers. (b) HM-HT fibers. (Sources: Morton and Hearle 1993 and manufacturers' data.)

properties, for example, flame resistance, antimicrobial action, and ultraviolet protection.

Most textile fibers will stand temperatures up to 200 °C, but at some higher temperature they melt, char, or chemically decompose in other ways. The thermoplastic fibers can be "permanently" heat set at temperatures appreciably below their melting points, and temporarily set at lower-temperature transitions. Similar temporary setting in cellulose and protein fibers is strongly influenced by wetting and drying.

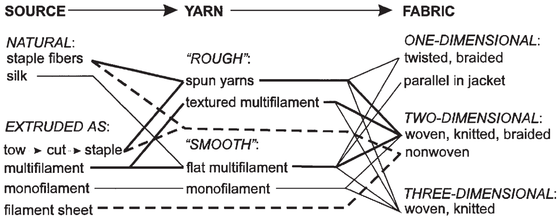

SOURCE
YARN
FABRIC
NATURAL:
staple fibers
silk
EXTRUDED AS:
tow ► cut ► staple
multifilament
monofilament
filament sheet
"ROUGH":
spun yarns
textured multifilament
"SMOOTH":
flat multifilament
monofilament
ONE-DIMENSIONAL:
twisted, braided
parallel in jacket
TWO-DIMENSIONAL:
woven, knitted, braided
nonwoven
THREE-DIMENSIONAL:
woven, knitted

Figure 2
Principal routes for assembly of fibers into fabrics.

1.5 Textile Processes and Products

Figure 2 outlines the major ways in which textile fibers are assembled into fabrics, which then have to be fabricated into the wide range of final products. In order to give yarns any strength, the natural staple fibers have to be twisted, entangled or bonded together. This is not necessary for continuous filaments, though limited interlacing or twisting is used to give some cohesion to the yarns. In order to get a rougher "natural" fabric character, coarse tows are made and cut into staple fibers for spinning, often in blends with natural fibers. Texturing processes can be used to set or trap filaments in helical or looped paths, in order to give more stretch and bulk to continuous filament yarns.

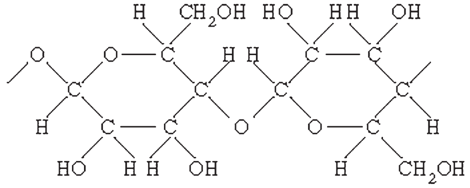
Most textiles are planar sheet materials. In woven fabrics, warp (lengthwise) yarns pass under and over weft (fill, crosswise) yarns. In braids, interlacing occurs between diagonal sets of yarns. In knits, yarns are looped together in neighboring rows, either across the length in weft-knitting or along the length in warp-knitting. Nonwoven fabrics are more-or-less random sheets of fibers, which are held together by adhesive bonding, entanglement, or stitching. One-dimensional assemblies are used in cords and ropes. Three-dimensional fabrics are used as composite preforms, either in the form of shaped sheets or thick structures, but also have other applications, such as, respectively, knitted garments and conveyor belts.

Dyeing of fibers, yarns, or fabrics in the wide range of needed colors involves selection of dyestuffs to match the chemical composition and structural accessibility of fibers. Acid, basic, and other reactive dyes are readily taken up by cellulose, protein, and nylon fibers, which are swollen in water. *Mordanting* fibers with metal salts allow dyes that form organometallic complexes to be used. Vat dyes are insoluble in water, but become soluble on reduction; fabric is reoxidized after dyeing. Disperse dyes, which are sparingly soluble in water, are used for thermoplastic fibers. Chemical and mechanical finishing of fabrics is used to stabilize dimensions, modify surfaces, and impart special functions.

2. Cellulose and Related Fibers

2.1 Basic Chemistry, Formation, and Properties

Cellulose, which is the main structural material in plants, is a condensation polymer formed by biosynthesis from carbon dioxide and water with glucose as the intermediate monomer. The chemical formula is



In terms of physical fiber structure and properties, the important molecular features are shown in Fig. 3. The long chain, which has a natural degree of polymerization (DP) of over 10 000, has direction. In the crystal form, cellulose I, in natural fibers, the chains are parallel, but mercerization (see Sect. 2.2) or solution and regeneration converts this form to cellulose II with antiparallel chains. The molecule is ribbon-like, which allows for easy twisting and relatively easy bending in one plane. The hydroxyl groups, which stick out of the chain, form hydrogen bonds within or between chains and with absorbed water.

Biosynthesis occurs at enzyme complexes within plant cells. At each complex, glucose units add on to about 30 growing molecules, which automatically crystallize as fibrils with a width of 3.5 nm. In this sense, natural cellulose fibers can be regarded as 100% crystalline. The lower fiber density, absorption of water, x-ray diffraction, and chemical evidence, which suggest the presence of around one-third disordered material, can be attributed to imperfect packing of fine fibrils with high surface area. Under genetic control, the fibrils are laid down in helical arrays, specific to particular plants, to form cell walls.

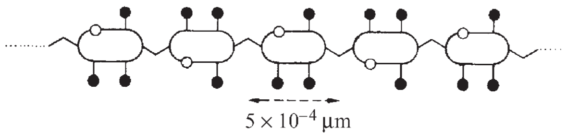


Figure 3
Essential features of the cellulose molecule.

The processing of cellulose to make regenerated fibers reduces the DP to between 250 in cheaper forms and 600 in improved forms. The amount of disordered material increases to about two-thirds, which doubles the water absorption.

Studies of the dynamic mechanical properties of viscose rayon, which are also generally applicable to other cellulose fibers, show glass-to-rubber transitions above 200 °C when dry and below 0 °C when wet. Consequently, wetting and drying leads to formation of creases. These can be reduced by crosslinking treatments with formaldehyde resins, which have the negative effect of making the fibers more brittle and less resistant to wear. Swelling in water makes for easy dyeing and chemical finishing of cellulose fibers (see also *Cellulose: Chemistry and Technology*#).

2.2 Cotton

Cotton accounts for almost half the world's fiber usage. Until challenged in price by polyester, its combination of low cost and useful properties made it the commodity fiber, which could be used in every application, including demanding technical textiles. Now, it is used more selectively, particularly for its comfort and its aesthetic and environmental appeal in clothing and household textiles.

The cotton fiber is a single plant cell, which is the seed hair of the genus *Gossypium*. Plant breeding has led to many varieties from different natural species. Genetic engineering is now introducing new traits. Many fibers grow from each of many seeds within the cotton boll, which forms after flowering. When the fibers are mature, the boll opens, and the fibers dry, ready for picking. Early opening, due to frost or disease, leads to thin-walled immature fibers. Cell formation starts with a primary wall, which grows to the full fiber dimensions in a few days. In the next 24–30 days, the secondary wall, which constitutes the main part of the fiber, is laid down in layers inside the primary wall with a helical angle of about 21°. At intervals, the helix reverses sense from left-handed to right-handed. At maturity, an open lumen remains at the center of the circular fiber. Drying causes the fiber to collapse into a flatter, bean-shaped cross-section. The helix reversals lead to the fiber taking the form of a convoluted ribbon. Natural waxes are present on the surface of the fiber. Raw fiber is slightly yellowish, but becomes white after bleaching. Some special varieties contain green or brown pigments. Cotton fiber dimensions range from short, coarse fibers (1.5 cm, 3 dtex) to long, fine fibers (5 cm, 1 dtex), with the bulk of the crop around 2.5 cm, 2dtex.

The stress–strain curve of cotton, as shown in Fig. 1, is slightly concave upward, and can be explained by the elastic modulus of the crystal lattice, modified by the effects of the helix angle, twisting at helix reversals, and removal of convolutions (Hearle and Sparrow 1979). The elastic recovery progressively decreases as extension increases, reaching a value of about 0.4 near break in dry cotton.

In the dry state, hydrogen bonding between fibrils resists shear, which occurs when a helical assembly is extended. The absorption of water introduces mobility between fibrils, and leads to a reduction in modulus and increase in break extension as humidity increases. Cotton, and other plant fibers, are unusual in having a higher strength wet than dry, due to the relief of internal stresses. The form of break also changes (Hearle *et al.* 1998). The tenacity and break extension of the weaker Indian cottons, which have been largely replaced by American varieties, have values of $0.19\,\mathrm{Ntex}^{-1}$ and 5.6%, compared to $0.45\,\mathrm{Ntex}^{-1}$ and 6.8% for strong Sea Island cottons. Typical values for the bulk of the crop are $0.35\ \mathrm{N\ tex}^{-1}$ and 7%.

Mercerization is a treatment of cotton with a sodium hydroxide solution. This disrupts the crystals, swelling and plasticizing the fibers but leaving sufficient cohesion to retain fiber identity and allow processing. The effects are reversed, though with changes in structure, when the caustic soda is washed out. If shrinkage is allowed, mercerized cotton will be weaker and more extensible, but stretching under tension in the process increases fiber strength and reduces extensibility. A main reason for mercerization is that it converts the fibers to unconvoluted round cross-sections with a smaller central lumen, thus increasing the fiber luster.

2.3 Other Natural Plant Fibers

Wood provides the commonest plant fibers, and is the raw material for regenerated cellulose fibers, but wood fibers are too short for textile processing (unless paper is regarded as a wet-laid nonwoven textile (see *Paper: History of Development*#; *Pulp and Paper: Wood Sources*#). *Kapok*, which is a hollow fiber used for buoyancy, is a seed fiber, and the coarse fiber *coir*, from the outer coat of the coconut, is a fruit fiber, but both have limited textile usage. Apart from cotton, the main textile plant fibers are extracted from stems (*bast fibers*) (e.g. flax, sunn hemp, true hemp, jute, kenaf, ramie) or stiff leaves (*leaf fibers*) (e.g. abaca (Manila hemp), sisal, henequen, phormium tenax (New Zealand flax)). At different times and locations, fibers from many other plants have been used. Currently, there is interest in promoting the use of waste products, such as pineapple fiber. The bast fibers were once of major importance, particularly before the supply of cotton from America, for clothing and other textiles; and the leaf fibers were dominant in ropes and cordage. However, with the additional competition from manufactured fibers, their use is now limited. The lower-cost fibers have inferior properties, and the superior fibers are expensive to produce.

Bast and leaf fibers are multicellular. The fiber ultimates are small and short, but in bundles they provide the reinforcement to stems and leaves. Extraction of the fibers starts with *retting*, which is a biological or chemical attack on the non-fibrous tissues, and is followed by mechanical beating and further textile processing. The size of the fibers, as they are used in yarns and fabrics, depends on the severity of the treatments. The helix angle in these fibers is lower than in cotton, typically around 10°, so that when, as in flax and ramie, they are almost pure cellulose, they have greater strength and lower break extensions than cotton. Other fibers, such as jute, contain substantial fractions of noncellulosic material, lignins and hemicelluloses, which lower strength and break extension.

Flax, which is used to make linen, and *ramie* are high-quality fibers with good color and an attractive appearance and handle in fabrics. *True hemp* (*Cannabis sativa*) was a major European and American fiber, which was largely displaced, though attempts are being made to revive its production. *Jute* was widely used for low-grade uses, such as sacking and carpet backing, but, in developed markets, has lost out to polypropylene. Trials for other uses are being made in jute-producing countries. In cordage, *abaca* and *sisal* were dominant, but have been replaced by synthetic fibers.

2.4 Regenerated Celluloses

The first factory to manufacture textile fibers, by extruding and regenerating cellulose nitrate (nitrocellulose) dissolved in ether and alcohol, opened in 1890 but this method is no longer used. The second process continues in limited production in Japan. Cellulose forms a soluble complex with cuprammonium hydroxide. The solution is extruded into a dilute acid bath, which precipitates the cellulose fibers. The third process, which in a variety of forms is now dominant, began commercial production in 1905 as *viscose rayon*. However, there are environmental problems with the viscose process, and it is doubtful if building a new, clean factory would be economically viable. This has led to research into organic solvents for cellulose, and semicommercial production of *lyocell* fibers began in 1988. These are likely to become more important. Other methods, such as making acetate yarns, stretching and then regenerating cellulose fibers, have been used at various times.

Marketing of cuprammonium and viscose rayon started with continuous filament yarns as a cheap “artificial silk.” Waste yarn was chopped into short lengths for processing into spun yarns on cotton and wool machinery. In the 1930s, high-tenacity yarns began to replace cotton in tire reinforcement, and the production of thick tows for conversion into staple fibers began. By 1950, staple had exceeded filament yarn production, and new variants with improved properties were introduced. Other developments and changing economics mean that rayon is no longer a low-cost commodity fiber, but is used to give added value either alone or in blends with polyester and cotton.

Viscose is an alkaline solution of sodium cellulose xanthate, which is made by reacting cellulose with sodium hydroxide and carbon disulfide. After filtering and aging, it is extruded into an acid bath. The filaments coagulate as regenerated cellulose fibers, which are stretched to increase strength and reduce extensibility. An early discovery was that the addition of zinc salts to the coagulating bath improved properties, by causing the regeneration to pass through an intermediate stage as zinc cellulose xanthate. Because hydrogen ions move faster than zinc ions, the zinc route occurred near the surface but the direct route operated in the center, giving a skin–core structure. Removal of excess solvent led to a collapse into a serrated cross-section.

The fine structure of "standard" viscose rayon is probably a classical fringed micelle structure, with little if any crystallographic chain folding. Chain molecules link crystalline micelles through tie-segments in amorphous regions. The crystals hold the structure together, deformation and rupture occurring in the disordered material. Chain ends between micelles reduce strength, but their effect can be minimized by increasing molecular weight, which is done for superior fibers, and by reducing the size of micelles and the intervening lengths. The last explains the influence of zinc: the skin structure has a finer texture than the core. For general textile purposes, moderate or low strength and a break extension around 20% was suitable, but for tire cords the stretch in manufacture was increased to give higher strength and reduced extension. The early tire cord yarns were skin-plus-core fibers, but the process was then modified to increase the skin thickness and eventually make stronger, "all-skin" fibers with a round cross-section. Similar techniques were later applied to more extensible staple fibers. By forcing the formation of an asymmetric skin, which contracts differentially, it is possible to make fibers that crimp spontaneously on drying.

In the 1950s, modifications, such as extrusion into a bath of a weak acid, made it possible to coagulate fibers of the cellulose derivative, stretch these, and then regenerate the fibers. This gives a fibrillar, instead of a micellar, fine structure. The resulting high-wet-modulus (HWM) fibers, which are also known as *polynosic* fibers, are closer to cotton in properties. Other developments included the production of hollow filaments, which give added bulk or can be collapsed in to ribbon forms, bulky multilobal fibers from shaped spinneret holes, deep-dyeing fibers, flame-retardant fibers, and superfine fibers (~ 0.5 dtex).

In the dry state, the tensile stress–strain curve of standard viscose rayon, as included in Fig. 1, shows an initial stiff region, with a yield at 1% or 2% extension to a lower-modulus section, which stiffens somewhat toward break. This is a micellar composite, so deformation occurs in the amorphous regions with only limited stiffening by the crystallites. Extension is initially restricted by the hydrogen bonds between cellulose molecules. Yield occurs when these bonds break. In the wet state, mobile water molecules replace the direct hydrogen bonds, and the low modulus yield region effectively starts at the origin. The fibrillar structure of the HWM fibers gives a parallel composite, in which the crystals stiffen the structure. This has some effect when dry, giving more cotton-like properties, and a major increase in wet modulus.

Lyocell fibers are made by dissolving cellulose in a concentrated amine oxide, *N*-methylmorpholine-*N*-oxide [7529–22–8] (NMMO), solution, and then extruding into a dilute NMMO solution, which precipitates the cellulose. With effective recycling, the process is environmentally and economically attractive. Due to their high DP and fibrillar structure, lyocell fibers, such as *Tencel*, have good properties. Fibrillation, namely the splitting off of fine fibrils from the fiber surface, occurs easily. This is advantageous in enabling fabrics to be finished to "peach skin" textures and in bonding hydro-entangled nonwovens, but, if not wanted, its extent can be controlled by process adjustments.

2.5 Cellulose Acetate

Treatment of cellulose with acetic acid and acetic anhydride converts it into cellulose acetate. Partial acetylation, with one in six hydroxyl groups remaining, gives *secondary acetate*, which is soluble in acetone and can be dry-spun into fibers. Due to their poor crystallinity, secondary acetate fibers are comparatively weak and extensible. Their low modulus means that fabrics have low bending stiffness, which is useful when soft, drapable materials are in fashion. Acetate fibers have low moisture absorption and are thermoplastic. The largest usage of acetate fibers is now in cigarette filters.

Complete acetylation gives cellulose triacetate, for which more expensive solvents are needed. Triacetate fibers, although weaker than nylon or polyester, have the same heat-setting advantages, and found a market when the synthetic fibers were expensive.

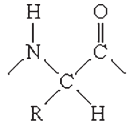
2.6 Alginate Fibers and Other Polysaccharides

Salts of alginic acid, derived from seaweed and with-COOH replacing the -CH_2OH of cellulose, can be made into fibers. Their main use is as calcium alginate fibers, which are soluble in hot, soapy water, and can be used to give temporary strength to a fabric. Calcium and sodium alginate fibers, and the other polysaccharides, chitin and chitosan, are used in wound dressings and other medical textiles.

3. Protein Fibers

3.1 Basic Protein Chemistry

The basic repeat unit in proteins is



The diversity of proteins is due to the choice and order of side groups -R, which derive from 20 amino acids. The side-groups are inert (-H and hydrocarbons), acidic, basic, hydroxyl-containing, and more complex. Of particular importance are two anomalous forms. Proline is a ring structure, which joins on to the main-chain carbon in place of -H and distorts chain geometry. Cysteine is formed as -CH_2SH, but can be oxidized to cystine, -CH_2SSCH_2-, which forms cross-links within or between chains. In addition to the cystine cross-links, covalent isopeptide bonds between lysine and glutamine, acid–base links, and hydrogen bonds influence fiber properties.

The residue sequence, which is controlled genetically, causes protein molecules to take up specific geometric forms. Some protein molecules, or parts of molecules, crystallize in one or other of two lattice forms. In variants of the α helix of simple polypeptides, the chains follow helical arrays, with internal hydrogen bonds and different chemical and geometrical repeats. Intermediate filaments (IFs) contain 32 chains, consisting of coiled-coil dimers, packed in tetramers and then octamers to form the microfibrils of wool. Alternatively, proteins crystallize as β sheets of extended molecules linked by hydrogen bonds between neighboring chains. Under tension, intermediate filaments undergo a reversible $\alpha \rightarrow \beta$ crystal structure transition, with an extension of 120% for an ideal α helix, but down to 80% in wool. Another class of proteins are globular, formed by individual chains folding into specific but irregular forms with internal cross-links.

3.2 Wool

Wool was formerly a more expensive commodity fiber, used for a wide variety of purposes. It is more resilient and durable than cotton, and can be made into warm, attractive clothing, upholstery, and carpets. Some of the attributes of wool have been replicated in synthetic fibers, and it must now compete on quality and tradition. Production costs are high, and farmers can change to other products. It has been demonstrated that genetic engineering can change the structure and properties of wool, though not yet in beneficial ways. This may be a route to a future for different wools with characters aimed at particular high-quality markets. Wool fibers are slightly elliptical, with mean diameters between 15 μm and 50 μm. Finer fibers command a higher price. A typical length of Australian Merino wools (~22 μm) is 100 mm and of coarse New Zealand wool (~35 μm), used in carpets from sheep grown mainly for meat, is 150 mm.

Wool, which is multicellular, grows under genetic control in a sequence of stages out of skin follicles at a rate of around 100 mm per year, and has the most complex structure of any textile fiber. Characteristic features are present at many levels. Chemically, there are families of keratin proteins of types I and II and of a variety of keratin-associated proteins (KAPs). The major part of the keratin I and II molecules pack into crystalline IFs, but there are also terminal groups (tails), which are rich in cysteine and project from the IFs. The KAPs are globular proteins, some of which are rich in cysteine while others are rich in glycine and tyrosine. The fine structure consists of a parallel assembly of IFs (microfibrils) with a diameter of about 7 nm, spaced about 10 nm apart in a matrix of KAPs. The IF structure, which is laid down first, is now well documented, but there is considerable uncertainty about the matrix, which is laid down later. The final actions are keratinization, which converts cysteine into cystine, and coating the fiber surface with a lipid layer. Some of the cystine cross-links will be within globular KAPs, but it is likely that some will be between globules and also to the tails of the IF proteins, which act as spacers between IFs. In this scenario, the matrix would act as a fairly strongly cross-linked rubber with a maximum extension of about 50%.

Most of the fiber consists of spindle-shaped cells held together by the cell membrane complex (CMC), which is rich in lipids. Cells in this central cortex are of two types, separated by a boundary, which rotates along the fiber. In the *para*-cortex the IFs are oriented parallel to the fiber axis; in the *ortho*-cortex the IFs are associated in helical arrays as macrofibrils. On the surface there are several layers of a cuticle, which contains scale cells, overlapping like tiles on a roof.

The structural features determine the specific properties of wool. The surface scales cause a directional frictional effect, which leads to *felting* (fiber entanglement by wet agitation) as movement is easier with scales than against scales. Felting is a defect, which causes shrinkage and can be avoided by chemical removal of scales, but can also be used to make fabric. Differential shrinkage of *ortho*- and *para*-cortex on drying causes fibers to bend, and the rotation of the boundary converts this into helical crimp, which gives bulk to wool yarns. The CMC generally holds the cellular composite together, but if it is weak it may allow slip and reduce fiber strength. The many amino acid side-groups present allow for a range of chemical treatments, including the introduction of covalent cross-links to give crease-resistance.

The tensile stress–strain curve of wet wool is stiff up to 2% extension, but then yields with an almost constant stress up to 30% extension, followed by a sharp change to a stiffer postyield region up to break at about 50% extension. A feature, which is unusual for yielding materials, is that recovery is complete from 30% extension or less, though on a curve at a lower stress level. Above 30% there is a small loss of recovery. In the dry state the yield stress increases, and recovery is incomplete. However, subsequent wetting will restore the fiber to its initial state. This is an important factor in the durability and appearance retention of wool fabrics.

The tensile properties have been explained (Chapman 1969) as the mechanics of a composite system. Up to 2% there is deformation of the IF crystal lattice with only a small contribution from the elastomeric matrix. At 2%, the $\alpha \rightarrow \beta$ transition is initiated in IFs and, locally, the IF stress drops from the critical value for nucleation of the β form to the equilibrium value for the $\alpha \rightarrow \beta$ transition, with stress transfer to the neighboring matrix, which extends to 30%. Only a limited zone can open, but the sequence proceeds with more zones opening until all the zones have joined up. Beyond 30% extension the matrix is stretched more, and the elastomer stress–strain curve is followed. Some cross-links break, which accounts for the loss of recovery and leads to fiber rupture in a granular form typical of a composite structure. Nucleation is not needed in recovery, which, below 30%, follows the elastomer curve with a shrinkage of opened zones until they disappear and the extension curve is rejoined below 1% extension. In the dry state, hydrogen bonding causes the matrix to have an initial stiff region, before yielding. The evidence is strong for the above explanation, but alternative views are still current (Feughelman 1997).

3.3 Other Hair Fibers

The hairs of other animals (and humans) have structures that are generally similar to wool but have specific differences. Coarse hairs, such as cow or horse, were used for undemanding uses, such as furniture stuffing, but also for special purposes, such as violin bows, where long fibers are needed. Other hairs, which find textile uses, are finer than wool or have other characteristics, which give particular qualities to fabrics. The main sources are varieties of goat, such as mohair and cashmere, camel, alpaca, vicuna, and angora rabbit.

3.4 Silk and Spider Silks

Silks differ from other natural fibers in not being formed as living cells, being extrusions of protein solutions which coagulate on drying. Silk used in the textile industry comes from the covering of the cocoon of the silk moth *Bombyx mori*. Pairs of triangular filaments (2–3 dtex) are stuck together by gum, and can be unreeled after soaking the cocoon in hot water. Each cocoon yields about 700 m of useful length of silk. As the only continuous filament yarn available to the textile industry until the twentieth century, silk gave expensive, lustrous fabrics. Even today, the shape of the fibers and their bulk and surface properties give special qualities, which command a high price in luxury fabrics. Production involves feeding mulberry leaves to silk worms, which then change to larvae within cocoons. Most cocoons are harvested for filament production. The silk from larvae that hatch out to continue the cycle is chopped into staple fiber for spun silk fabrics. There is also some use of wild tussah silk.

Silk is a natural block copolymer. Segments of a simple repeating structure of glycine, alanine, serine, and tyrosine form β-crystalline regions linked by segments with more diverse chemistry. Cysteine is absent. Silk is the toughest (highest work of rupture and good recovery) of natural textile fibers, due to moderately high tenacity and break extension. This led to uses such as parachutes until it was superseded by nylon.

Spider silks have not found textile uses, but have remarkable and variable properties. Different proteins are formed from the same glands for different purposes. Drag-line silk is extremely strong, and derives from an oriented liquid crystal solution, similar to production of HM-HT fibers such as *Kevlar*. Other spider silks in the web are more extensible and elastic. Currently, there is interest in the genetic engineering of spider silk proteins by bacteria as a route to polymers for high-performance fibers from aqueous solutions, but other factors such as the form of the "spinneret" in the spider and the relatively slow extrusion rate may be critical (see *Silk Produced by Engineered Bacteria*[#]).

3.5 Regenerated Protein Fibers

Fibers can be regenerated from solutions of proteins, but the price/property ratio was not sufficiently favorable for continued commercial production. The fibers are weaker than wool and have poor elastic recovery. The most successful was made of casein from milk, but proteins from corn and ground-nuts were also used.

4. Melt-spun Synthetic Fibers

4.1 General Features: Polyamides and Polyesters

The invention by Carothers of the polyamide *nylon 66*, which was commercialized by DuPont in 1938, marks the real start of a major synthetic fiber manufacturing industry. It was followed by *nylon 6*, invented in Germany, and by the polyester *polyethylene terephthalate* (PET), which has now overtaken nylon in world production. These are all polymers with melting points between 215 °C and 260 °C. Molten polymer is extruded, cooled, and solidified, and stretched (drawn) to give orientation. Continuous filament yarns contain 1 to about 1000 filaments. Staple fibers are cut from massive tows.

In the first production method, polymer chips were melted, extruded and wound up at under $1000\,\mathrm{m\,min^{-1}}$ as unoriented yarns, which were partially crystalline for nylon but amorphous for polyester. In a second stage, the undrawn yarn was stretched by draw ratios of around × 4. Later, combined spin–draw processes were introduced, with first-stage rollers at around $1000\,\mathrm{m\,min^{-1}}$ and wind-up at $3000\,\mathrm{m\,min^{-1}}$. A major change, particularly important in polyester, was the production of *partially oriented yarn* (POY) by direct wind-up at high speed. The orientation results from the rapid elongation and attenuation of the molten threadline before solidification. The necessary residual draw ratio of around × 1.5 is imparted simultaneously with the following textile operations: yarn texturing or warp-drawing. Even higher-speed winding machinery, 5000–$10\,000\,\mathrm{m\,min^{-1}}$, allows production of comparatively low-strength, high-extension yarns suitable for direct use in fabrics for some purposes. Continuous polymerization, which feeds molten polymer directly to extrusion, is now common.

Other heat treatments, with tension or relaxation, can be applied to change the tensile, thermal shrinkage, and other properties. For general textile purposes, fibers of moderate strength and break extensions around 20% are produced, but for tire cords and other technical uses higher strengths and lower break extensions (~10%) are used. Fiber cross-sections are usually circular, but other shapes, such as bulky trilobal fibers, can be formed from shaped spinneret holes. Finishes are applied to reduce electrostatic charging and to give frictional properties needed for different markets. Internal additives, such as delustrants, pigments, or conducting carbon or polymer, can be incorporated. For some uses, bicomponent fibers, with skin–core or side-by-side forms, or multicomponent fibers are produced. An important example of the latter is the islands-in-a-sea process which gives very fine microfibers embedded in a matrix, which can be removed.

Nylon 6 and 66 and PET are characterized by polymers with long repeat units that contain flexible inert sequences alternating with more interactive sequences, which in PET are also stiffer. This gives rise to thermal transitions which have led these polymers to dominate the market for synthetic fibers. At some temperature around −100 °C, the amorphous regions of the polymer change from a fully glassy state to give some extensibility as freedom of rotation around C–C or similar bonds becomes possible. At around 100 °C, another second-order transition occurs, as hydrogen bond mobility develops in nylon and benzene ring interactions become free in PET. In addition to lowering the modulus, this opens up the structure for dyeing. At about 80 °C below the melting point the structure loosens up still more, probably by some form of crystal mobility, but possibly in PET by transesterification in amorphous tie molecules.

Heat setting, namely the fixing of fibers within yarns and fabrics in new shapes, is an important attribute of nylon and polyester fibers. The intermediate transition at around 100 °C causes a temporary set, which can be released by reheating. The higher transition allows "permanent" heat-setting treatments to be applied to nylon 66 and PET yarns and fabrics between 180 °C and 240 °C, when fibers start to stick together. The temperatures are lower for nylon 6. Setting processes are used in the production of textured yarns and to fix creases or otherwise stabilize fabrics. In nylon, it is necessary to go to more severe conditions to overcome a prior treatment, but PET can be reset at any temperature in the range.

Superficially, the extrusion, attenuation, and solidification of a molten threadline looks like a steady process, and the macroscopic relations between forces, elongation, and rheology and between heat transfer and temperature have been extensively studied. However, at the molecular level the motions may be more complicated. The semicrystalline structure is established at high speeds (1ms corresponds to 5cm at $3000\,\mathrm{m\,min^{-1}}$). In addition to polymer dynamics in situations far from equilibrium, the phenomena will involve aspects of the emerging science of complexity, chaos, fractals, kinetics of nonhomogeneous processes, nonlinear irreversible thermodynamics, and, possibly, quantum superposition, but it is not easy to see how to apply these ideas.

Despite studies by many analytical techniques, there is still considerable uncertainty about the fine structure of these fibers. Degrees of orientation, which vary with stretch in processing, can be measured. Crystallinity can be estimated on the basis of a scale from crystal to amorphous of a parameter such as density, and is always close to 50% in fibers as used. Some indications of the size, spacing, and linkage of crystals can be obtained, but the details remain obscure. Partly this is due to the many ways in which molecules can be arranged in a partially oriented, semicrystalline structure, and partly to the fact that the structures formed and modified by annealing at temperatures close to the melting point will vary with processing history. There have not been comprehensive investigations and interpretation of well-characterized samples by the several available techniques. Almost all pictures drawn to illustrate views of structure have been simplistic and not quantitative.

The common view is that the fine structure is a modified fringed micelle form. Crystallites, with a mix of fringing and folding at their ends, are stacked

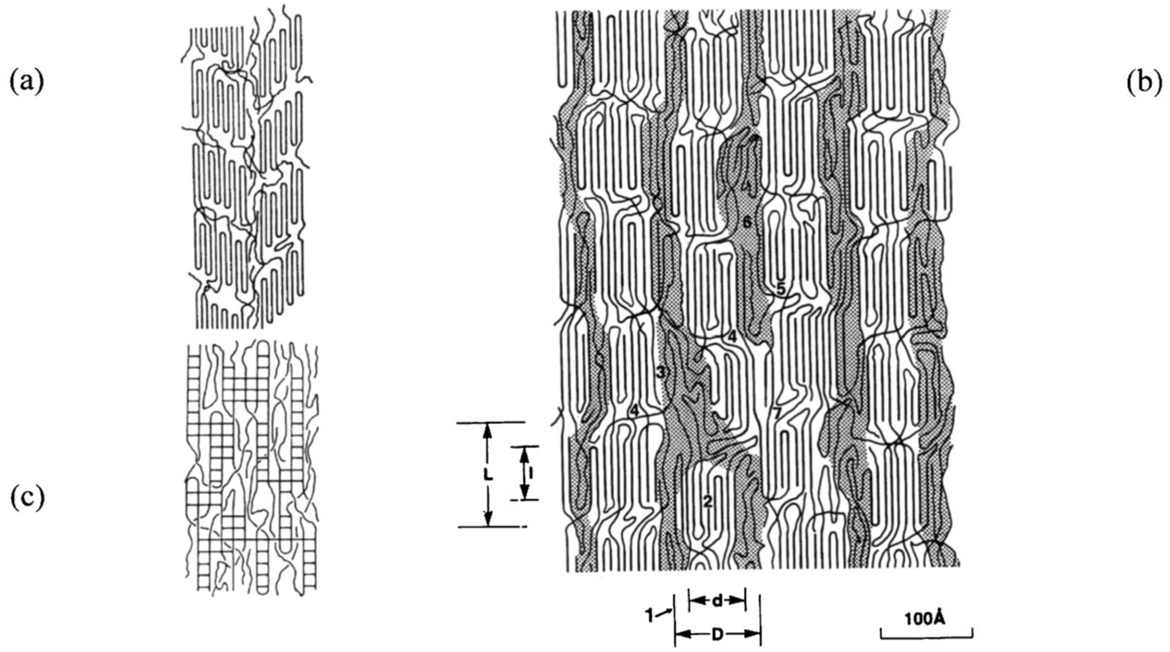


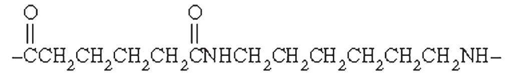
Figure 4
(a) Schematic view of nylon fine structure, intended to explain the small-angle X-ray diffraction pattern by the angled lattice repeat. (b) Fine structure of nylon 6, as interpreted from SAXS and WAXD data (Murthy *et al.* 1990). (c) Alternative continuous partially ordered structure.

in a pseudofibrillar array and linked by tie molecules, as suggested in Fig. 4(a). Figure 4(b) is an attempt to be more specific on the structure of a nylon 6 fiber and, while retaining the main features of more idealized representations of the common view, introduces some variability and three-dimensionality. Other views have been suggested. The long repeat length means that defects, such as those that occur in polyethylene due to additional units, are not possible without disruption of crystals, but, in contrast to Figs. 4(a) and (b) and similar pictures, there could be defect forms in which molecules follow paths across crystallites and emerge from the sides. More rapid quenching may give continuous structures of partial order, such as illustrated in Fig. 4(c), in which the boundaries between crystalline and amorphous regions are less well defined.

The partial order and molecular characteristics give strong, extensible nylon and polyester fibers. Tensile rupture is a ductile break in which a crack opens to a V notch, before final transverse rupture. Apart from the range of possibilities for each fiber type, there will be generic differences between nylon 6, nylon 66, and PET. Cracks induced by shear stresses move across nylon fibers at an angle of about 10° but are parallel to the axis in polyester fibers, which suggests that crystallites may have higher aspect ratios. The structural differences, together with the stiffening effect and different intermolecular forces of the benzene ring in amorphous regions, lead to the significant differences between nylon and polyester in properties and uses.

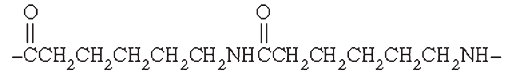
4.2 *Polyamides: Nylon 6 and 66*

The chemical formulae of nylon 66 and two repeats of nylon 6 are



$$-\overset{O}{\overset{\|}{C}}CH_2CH_2CH_2CH_2\overset{O}{\overset{\|}{C}}NHCH_2CH_2CH_2CH_2CH_2CH_2NH-$$

Nylon 66



$$-\overset{O}{\overset{\|}{C}}CH_2CH_2CH_2CH_2CH_2NH\overset{O}{\overset{\|}{C}}CH_2CH_2CH_2CH_2CH_2NH-$$

Two repeats of nylon 6

The same units are present with a slight rearrangement in two repeats of nylon 6 as in one repeat of nylon 66. The main difference is that the nylon 6 molecule has direction, which is absent due to the symmetry of nylon 66. This affects the possibilities for crystal packing, and is probably the reason for the lower melting point of nylon 6.

Although some sigmoidal curvature can be put into the tensile stress–strain curve of nylon fibers by temporary setting, the curves are usually almost linear up to a yield region close to the break stress. The modulus is comparatively low. Depending on the last treatment by the manufacturer, nylon fibers will often shrink by 10% in boiling water, which lowers the modulus. Alternatively, setting under tension may increase length by 10% and raise the modulus. Moderately high strength and break extension combine to make nylon a tough, durable fiber. Elastic recovery is good, being around 70% from near break stress, but time-dependent. Nylon shows an appreciable amount of primary (recoverable) creep and stress relaxation. Due to the presence of -CONH- groups, nylon absorbs water, which leads to changes in properties. The glass transition drops from near 100 °C when dry to near 0 °C when wet, and strength is reduced. Near the transition temperature the dissipation factor (tan δ) due to internal friction, increases to over 0.1, and above the transition the modulus is lower.

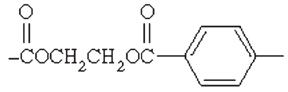
The first uses of nylon in parachute fabrics and ladies' stockings resulted from its toughness and durability. After 1945, markets developed for nylon as the first "ease-of-care," "wash-and-wear" fabric, due to the ability to give a permanent set. Setting also led to high-stretch yarns, which were made by twisting, heat setting, and untwisting, at first in separate steps and then by continuous false twisting. Nylon production was overtaken by polyester in the 1970s, and nylon is now used selectively for applications for which its properties are well suited. These features are low modulus, high work of rupture, good durability when dry, and resilient recovery from large deformations. In addition to underwear and some other items of clothing, there is a large market for carpet yarns, which can be made either from staple nylon spun in the same way as wool or from bulked continuous filament yarns (BCF) made by a jet-screen buckling and setting process. Industrial uses, such as ropes, come when a limited number of applications of high forces, for example in the shock loading of climbing ropes, are expected. When wet, nylon has poor wear resistance in applications such as marine ropes, due to shear cracks from surface abrasion under repeated cycling at moderate loads.

4.3 Other Nylons

Polyamides with more -CH_2- groups, such as nylon 11 or nylon 610, will have lower melting points, density and moisture absorption, and will tend to be more extensible. They have mainly found application in thick monofilaments as bristles. Reducing the number of -CH_2- groups, for example in nylon 4, will give a stiffer fiber, which is advantageous for some applications. Attempts have been made to follow this route, but there is a large barrier of lower price in the larger volume of nylon 6 and 66 to overcome. In the 1970s, DuPont made a more expensive, luxury fiber, *QIANA*, from the polymer of an amine with two cyclic rings, *trans*,*trans*-di(4-aminocyclohexyl)methane, with dibasic acids having 6–10 -CH_2- groups; but sales did not justify continuing production.

4.4 Polyethylene Terephthalate

PET (or 2GT) is the ester of ethylene glycol and terephthalic acid, which are polymerized either directly or through other intermediates, and has the chemical formula



The -COCCOC- sequence acts in a similar way to the -CH_2- sequence in nylon. The benzene rings make the molecule stiffer, and form intermolecular cross-links analogous to hydrogen bonds.

PET has a good combination of properties for a wide range of textile products. Depending on the molecular orientation, combinations of moderate strength and high break extension or high strength and moderate break extension give a high work of rupture. The initial modulus is about three times that of nylon, which is beneficial in many applications. Elastic recovery from high loads is less than for nylon, but from strains under tensions up to about half the break load is better. The extent of creep is small. The glass transition is above 100 °C and, since PET absorbs little water, is almost the same wet and dry. This means that creasing is not imposed by normal washing and drying. "Permanent" heat setting can be achieved repetitively above 180 °C. Twist-textured PET yarns are processed continuously over two heaters. The first stage of twist–set–untwist gives a yarn with a potential for high contraction, like the nylon stretch yarns, but second-stage setting after contraction by around 15% gives a yarn with high bulk but limited stretch.

The majority of polyester filament yarns are produced as POY yarns at around $3000\,\mathrm{m\,min^{-1}}$. This speed induces an incipient crystallinity, which prevents the rapid structural aging of yarns spun at lower speeds. Lower-strength yarns, for undemanding end uses in flat-filament fabrics, can be produced directly at speeds from $5000\,\mathrm{m\,min^{-1}}$ to $10\,000\,\mathrm{m\,min^{-1}}$. Staple fiber is cut from large tows, which have been hot-stretched in wet conditions and crimped by passing through stuffer boxes. Continuous filament webs are extruded and thermally bonded as nonwoven fabrics.

However, there is considerable diversity beyond the simplest round cross-section fibers of PET homopolymer, which are used in the commodity apparel textile market. Microfibers and shaped fibers give special effects in handle, drape and other aesthetic characteristics, particularly in the *Shingosen* fabrics from Japan. Examples are the peach-skin feel from microfibers and the characteristic luster and rustle of silk from three-petal shaped fibers. Copolymers are used to enhance dyeability, or to give differential dyeability by incorporating groups attractive to acidic or basic dyes. Pilling, which can be a nuisance, is reduced by using weaker fibers, which allow pills to break off.

Industrial grades are processed to give strengths that are the highest of any except the newer high-performance fibers, and special finishes are applied to meet needs such as promotion of adhesion to rubber in tire cords or low-friction to reduce internal abrasion in marine rope grades.

Owing to the ease of manufacture, large volume, and good properties, PET polyester has become the low-cost fiber for a universal range of purposes. This can be shown by the list below of chapters in the applications section of a book celebrating 50 years of polyester (Brunnschweiler and Hearle 1993). Each chapter covers many products. The last item in the list is typical in going from cheap disposable nonwoven surgeons' gowns and sanitary products to expensive knitted tubes for arterial grafts and braided replacements for ligaments.

Woven and knitted apparel fabrics. Man-made suedes and leathers. Household products. Fashion and function in Italy. Textiles for the home. The revolution in sewing threads. Competing with feathers and hair: the *Fiberfill* story. Tyres and industrial textiles. Polyester inside cars. Forming paper: a success for monofilaments. Ropes and nets. Building with polyester fabrics. Geotextiles. Composites: a lost opportunity or an exciting future. Medical textiles: perfect fit for a critical use.

Once five times the price of cotton, the PET commodity fibers and yarns are now similar or lower in price and account for about 25% of world fiber production. The use of the cheapest PET fibers in lower-quality fabrics has given polyester a poor image. This is not justified when premium fibers are used in quality fabrics. It is possible to combine excellent aesthetics with durability and ease-of-care.

4.5 Other Polyesters

Analogues of PET (2GT), derived from other glycols with more -CH_2- groups in the chain, give fibers with broadly similar properties. Their viability depends largely on economic factors. There has been commercial production of 4GT. Genetic engineering of bacteria to convert sugar from corn into trimethylene glycol opens a route to 3GT.

Polyethylene naphthalate (PEN) fibers, produced in experimental quantities in the 1990s, are made from the double-ringed naphthalene analogue of terephthalic acid. This development results from the commercialization of an economic route to intermediate chemical production. The fibers have a higher modulus and a higher melting point than PET, which is useful for some industrial applications.

4.6 Polyolefins

The hydrocarbon polyolefins, *polyethylene* (PE) and *polypropylene* (PP), have the chemical formulae

$$\underset{\text{PE}}{-CH_2-} \qquad \underset{\text{PP}}{-CH_2CH(CH_3)-}$$

These are short repeat units without interactive groups. The polymers can be melt spun into fibers, but, except for toughness, strength, and break extension, do not share the similarities of nylon and polyester. Their density and melting points are lower; recovery is poorer; fiber feel is waxy; moisture absorption is zero; chemically they are inert and resistant to solvents; dyeing is impossible without modification; and coloration has to be achieved by adding pigments to the melt. The single large glass transition in PE makes it undesirably soft for most textile purposes, and the approach to melting cause its properties to fall off below 100 °C. The single transition in PP spans room temperature, which makes the fiber stiffer, but gives large energy losses in deformation.

PE has not penetrated the general textile fiber market. Thick monofilaments are used as an alternative to natural fibers in canvas fabrics and in ropes. *Tyvek*, which can be regarded as a stiff nonwoven textile or a tough paper, is flash spun from solution to give a fibrillated web, which is bonded by heat. Gel spun HMPE fibers are described in the high-performance fiber section below.

PP is more widely used, but 40% of its textile use is in the form of slit film, which is used for coarse fabrics such as carpet backing; there are similar uses for *Typar*, which is a directly produced, thermally bonded, continuous-filament web. The hydrophobic character of PP has been promoted as advantageous in thermal underwear. For nonwovens, its ability to be thermally bonded without undue compression of fiber webs is superior to polyester. However, PP continuous-filament yarns and staple fibers have mostly found textile markets where lower cost offsets poorer properties. Major uses are disposable nonwovens, and cheaper carpets, upholstery, and ropes. PP yarns are easy to produce, so that ropemakers, for example, have their own production facilities.

5. *Solution-spun Synthetic Fibers*

5.1 *Polyacrylonitrile (Acrylic)*

Polyacrylonitrile (PAN) was the third major synthetic fiber of the 1950s, but has not achieved the wide usage of nylon and polyester. Apart from providing precursor yarns for carbon fibers, PAN is almost entirely produced as staple fiber. The major uses are in warm, bulky fabrics, such as knitwear, blankets, and pile fabrics, as an alternative to wool. Its synthetic-fiber carpet market has been lost to nylon.

The basic chemical formula of PAN is

$$-CH_2CH(CN)-$$

In order to make dyeing possible, acrylic fibers (specified as $>85\%$ PAN) are copolymers with added reactive groups. Acrylic fibers can be melt spun, but are usually produced from solution by wet or dry spinning. The diversity of materials and processes leads to differences in properties in fibers from different manufacturers and in different options.

Acrylic fibers are weaker than nylon or polyester, yield at 2% extension, and have poor recovery from higher extensions. Breaks are granular. They are regarded as quasicrystalline, and have a major second-order transition near 100 °C. Fibers which have been highly stretched, or broken, show a high shrinkage when heated. Blending of no-shrink or preshrunk and high-shrink fibers is a way of making bulky yarns.

5.2 *Other Solution-spun Fibers*

Modacrylic fibers contain less PAN and more copolymer groups. Properties are generally similar to those of acrylic fibers, but the copolymer is used to give special properties such as flame resistance.

Vinyl and vinylidene polymers and copolymers were the first polymers to be used as fibers, but have not achieved widespread use, as can be seen from the small figure in Table 2 for "others," which also includes some high-performance fibers. *Polyvinyl chloride* (PVC) fibers are produced in the same way as acrylic fibers, and have somewhat similar properties. Polyvinyl alcohol (PVA) fibers have good properties, and have been produced in substantial quantities in Japan and China, but the cost/performance ratio has not worked out as favorably as some have expected.

6. *High-performance and Specialty Synthetic Fibers*

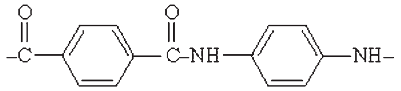
6.1 *General Features*

Following the success of nylon, polyester, and acrylic fibers, research moved to a second generation of high-performance, synthetic-polymer fibers. The major drive was for HM-HT fibers for industrial uses such as tire cords and for advanced composites and other defense-related applications. The structural requirement is to have long polymer chains (high molecular weight), fully extended and oriented parallel to the fiber axis. They will usually, but not necessarily, be highly crystalline. Two routes were found. The first uses stiff, interactive polymers, which form liquid crystals in solution or melt. These become oriented during extrusion. From solution, the technique is dry-jet, wet-bath spinning, with a gap between the end of the spinneret and the coagulating bath. The second route uses flexible, inert polymers, which form an orientable gel in concentrated solution. Unfortunately, no method has been found to produce the right structure in nylon or PET fibers, which would have technical and economic advantage.

Another need was to produce fibers with high temperature resistance. Glass, carbon, and ceramic fibers are high-modulus and high-temperature fibers, but, as mentioned in the introduction, are not the concern of this article. Another group of advanced polymer fibers has special physical or chemical properties.

6.2 *Para-aramid Fibers*

The aramid fiber *Kevlar*, *polyphenylene terephthalamide* (PPTA), was the first successful HM-HT polymer fiber. It is a *para*-aramid, with benzene rings from dibasic acid and diamine joined along the chain in 1,4 positions. The chemical formula is



The polymer cannot be melted, but degrades chemically at around 500 °C. To be spun, it is dissolved in concentrated sulfuric acid and extruded into an alkali bath. The fiber, as first produced, is highly crystalline and highly, but not perfectly, oriented. The structure consists of pleated sheets, which are slightly off-axis. This structure causes the stress–strain curve to be concave upward, and results in some creep under low loads. Further heat treatments under tension increase orientation and give higher modulus forms. Other techniques can reduce modulus, as needed in some applications.

Covalent bonds give the high axial properties, including a tensile strength around 3 GPa, but intermolecular forces are weak. Hydrogen bonding between -CONH- groups occurs only in one plane. Tensile rupture results from axial splits, which run across the fiber, giving long fibrillar forms. The shear modulus is low. Axial yield occurs at about 0.4 GPa. Although the low compressive strength is a disadvantage in many ways, it has the advantage that the fibers are not brittle in bending. Yielding on the inside of a bend allows fibers to bend back on themselves through 180° without breaking. Kink bands can be seen internally with polarized light microscopy and externally with electron microscopy. Although a few repetitions of sharp bending do not cause major damage, a few thousand cycles will cause breakage by flexural fatigue. This is seen as axial compression fatigue when component yarns in ropes are allowed to go into compression. The nature of the tensile break and the compressive yielding lead to high-energy absorption in structural impacts.

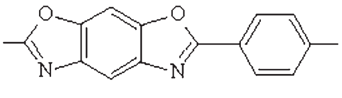
Technora is a copolymer aramid fiber in which half the *p*-phenylene diamine is substituted by 4,4′-diaminodiphenyl ether. Its properties are similar to PPTA fibers, but with differences which may be advantageous in some applications.

It was hoped that *para*-aramid fibers would find a large market in tire cords. However, apart from the high price, they are stronger than is needed for the general market. Going from cotton to rayon to nylon and polyester, it was possible to reduce the amount of cord needed as strength was increased. However, there is a limit to the sparseness of cords within the tire, so that advantage could not be taken of the added strength of aramids. Their use depends on applications where high strength and high stiffness justify a high price. This includes: heavy-duty tires; light, strong ropes and industrial fabrics; ballistic impact resistance; and advanced composites.

6.3 Other Liquid Crystal Fibers

The *para*-aramid polybenzamide (PBA), which is the analog of nylon 6, and other polymers with benzene rings in the chain are alternatives to PPTA, which is the nylon 66 analog, but have not been commercialized as of 2000. Superior properties come from polymers with more complicated rings, which were the subject of US Air Force research.

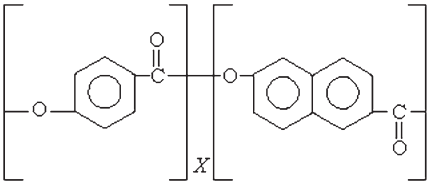
One of these, *polybenzoxazole* (PBO), has been commercialized by Toyobo as *Zylon*. The chemical formula is



Production methods, structure, and properties are similar to those of aramids, but the modulus and strength are higher.

Experimentally, *M 5* fibers with higher shear modulus and strength and higher compressive yield stress have been produced from a polymer with a similar ring structure to PBO, but with -OH side-groups giving hydrogen bonding in both planes. HM-HT cellulose fibers have been made from solutions in phosphoric acid.

Vectran is a melt-spun fiber, based on the fully aromatic polyester *polyphenylene terephthalamide*, which has the following chemical formula:



Limited copolymerization gives a thermotropic liquid-crystal material, which orients in flow through the spinneret. Melt spinning is much cheaper than solution spinning with strong acids, but this is offset by the need for an expensive subsequent heat treatment to build high molecular weight. The properties of *Vectran* are generally similar to aramids, but with specific differences, such as an absence of creep, which are beneficial in some applications.

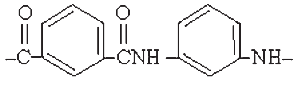
6.4 High-modulus Polyethylene

HMPE fibers are made by gel spinning of ultra-high molecular weight polyethylene. The extruded gels, which form in concentrated solutions, can be stretched to give highly oriented, chain-extended, highly crystalline fibers. The strength and modulus are higher than for aramids, but the fibers have the disadvantage of a low melting point. Properties start to fall off appreciably as temperatures rise above 40 °C. Substantial creep is due to molecules sliding over one another. The rate of creep increases with tension, and ultimately leads to creep rupture. Post-treatments can increase modulus and reduce creep. There are major differences in the incidence of creep among the different grades from different manufacturers. In one set of tests, one HMPE yarn broke after 4days at 30% of break load, whereas another lasted 123 days. Improved grades have been introduced since these tests were carried out. The converse of creep failure is the excellent resistance to high-speed loading. HMPE fibers have low compressive strength, but are somewhat more resistant to axial compression fatigue than aramids. Gel spun HMPE fibers find applications in ballistic protection, in high-performance ropes and industrial fabrics, and in some composites.

An alternative route to HMPE fibers is super drawing of melt-extruded polymer. However, the strength is about half that of gel-spun fibers, and commercial production has not been successful.

6.5 Thermally Resistant Fibers

The meta-aramid fiber *Nomex*, which is *poly*-m-*phenylene isophthalamide*, has the chemical formula



The 1,3 links between benzene rings gives a less compact molecule than the 1,4 *para*-aramids. The polymer can be spun into fibers with mechanical properties comparable to general textile fibers. The advantages of *Nomex* are excellent thermal stability, flame resistance, and good dielectric properties. It is used in industrial apparel, aircraft upholstery, composites, and electrical insulation subject to high temperature.

Other linear polymers, such as poly-*m*-phenylene benzimidazole and aromatic polyimides, can be used to make thermally resistant fibers. Flame-retardant fibers, with the FTC designation *novaloid*, are made from cross-linked phenol–formaldehyde polymers.

6.6 Elastomeric Fibers

Elastic threads can be made from natural or synthetic rubbers, but it is not possible to make good fine fibers. Spandex fibers, such as *Lycra*, are made from segmented polyurethanes from the reaction of a diisocyanate with polyethers or polyesters. Amorphous soft segments give the elastic extension, and crystalline hard segments cross-link the structure and hold it together. The fibers are relatively weak, though the true stress at break is high, and have elastic extensions over 400%. They are resistant to physical and chemical damage in use, and are incorporated in fabrics when high stretch is required.

6.7 Fibers with Special Functions

There are a variety of fibers used in specialized applications, because of particular chemical and physical properties. *Teflon*, *polytetrafluorethylene* (PTFE), which cannot be melted or dissolved, is spun into fibers by heat sintering a dispersion of particles. The fibers are used where low friction or high chemical resistance are needed. In addition to the use of Teflon and polyesters in arterial grafts, several new fibers are being used for medical purposes. For example, fibers of polyglycolic acid are used for sutures, because they are assimilated by the body after the limited period for which they are needed to function.

A great variety of smart fibers are now being explored, and some are in commercial use. They are mostly made by incorporating additives. By way of example, the range seen on a visit to one Japanese company in 1993 comprised: exothermic, electrically conducting yarns, which can be incorporated as resistance elements to control temperature in fabrics; antistatic fibers for clean rooms; polypeptides that absorb and disperse water; fibers that melt and bind; filaments that vary in thickness to match natural nonuniformity; fibers with zirconium carbide that warm by absorbing near infrared and emitting far-infrared radiation; chitin fibers for wound dressing; activated carbon fibers to absorb and desorb chemicals; amorphous metal fibers with high strength and special magnetic properties; and glass fibers with special coatings for printed circuits.

7. Conclusion

Textile fibers exist in great variety and are used for many applications in a diversity of structures. Much more could be written on their place in materials science and technology than can be included in this article. However, in another sense, the materials are limited. Table 2 shows that about 80% of the world's textile fibers are now based on cellulose or polyester, 18% on nylon, polyacrylonitrile and polypropylene, and 2% on proteins. Some vinyl polymers, which have not achieved wide market acceptance, and the newer high-performance and specialty fibers, which are expensive, account for about 0.5%. The diversity is mainly achieved by form, shape, size, process parameters, copolymerization, additives, and finishes.

See also:

Carbon Fibers*; Cellulose: Chemistry and Technology#; Ceramic Fibers from Polymer Precursors*; Fibrous Reinforcements for Composites: Overview*; CVD Monofilaments*; Glass Fibers*; Paper: History of Development#; Metallic Filaments*; Glass Optical Fibers#; Silk Produced by Engineered Bacteria#; Spun (Slurry and Sol–Gel) Ceramic Fibers*; Whiskers*; Pulp and Paper: Wood Sources#

Bibliography

Alexander P, Hudson R F 1963 *Wool: Its Chemistry and Physics*, 2nd edn (revised by Earland C). Chapman and Hall, London

Brunnschweiler D, Hearle J W S 1993 *Polyester: 50 Years of Achievement*. The Textile Institute, Manchester, UK

Bunsell A R (ed.) 1988 *Fibre Reinforcements for Composite Materials*. Elsevier, Amsterdam

Chapman B M 1969 A mechanical model for wool and other keratin fibers. *Textile Res. J.* **39**, 1102–9

Cook J G 1968, 4th edn. *Handbook of Textile Fibres*, Merrow, Watford, UK, Vols. 1 and 2

Feughelman M 1997 *Mechanical Properties and Structure of Alpha-Keratin Fibres*. UNSW Press, Sydney

Hearle J W S, Lomas B, Cooke W D 1998 *Atlas of Fibre Fracture and Damage to Textiles*, 2nd edn. Woodhead Publishing, Cambridge, UK

Hearle J W S, Sparrow J T 1979 Mechanics of the extension of cotton fibers. *J. Appl. Polym. Sci.* **24**, 1857–74

Hongu T, Phillips G O 1997 *New Fibers*. Woodhead Publishing, Cambridge, UK

Lewin M, Pearce E M (eds.) 1985 *Fiber Chemistry*. Dekker, New York

Lord E 1961 *The Characteristics of Raw Cotton*. The Textile Institute, Manchester, UK

McIntyre J E, Daniels P N (eds.) 1995 *Textile Terms and Definitions*, 10th edn. The Textile Institute, Manchester, UK

Mark H F, Atlas S M, Cernia E (eds.) *Man-made Fibers. Science and Technology*. Interscience, New York, Vols. 1, 2, 3
Mauersberger H R (ed.) 1954 *Matthews' Textile Fibers*, 6th edn. Wiley-Interscience, New York
Moncrieff R W 1975 *Man-made Fibres*, 6th edn. Newnes-Butterworth, London
Morton W E, Hearle J W S 1993 *Physical Properties of Textile Fibres*, 3rd edn. The Textile Institute, Manchester, UK
Murthy N S, Reimschuessel A C, Kramer V J 1990 Changes in void content and free volume in fibers during heat setting and their influence on dye diffusion and mechanical properties. *J. Appl. Polym. Sci.* **40**, 249–62
Nakajima T (ed.) 1994 *Advanced Fiber Spinning Technology*, English edition (eds. Kajiwara K, McIntyre J E). Woodhead Publishing Cambridge, UK
Reneker D H, Chun I 1996 Nanometer diameter fibers of polymer, produced by electrospinning. *Nanotechnology* **7**, 216–23
Ziabicki A, Kawai H (eds.) 1985 *High-speed Fiber Spinning*. Wiley-Interscience, New York

J. W. S. Hearle
UMIST, Manchester, UK

Textile Fibers: Mechanical Properties

Aside from their traditional use in clothing or home furnishing and decorating, textile fibers are widely employed in technical applications such as ropes, containers, covers, belts, filters, tires, and the expanding range of fiber-reinforced composites. While the strength of fibers is just one of several design criteria in the fabrication of, for example, clothing, it is generally the major and crucial factor in engineering applications. The resistance of fibers to failure under imposed mechanical stresses, given certain thermal and chemical influences as well as exposure to UV light, depends on chemical composition, but more importantly on supramolecular structure (morphology).

1. Fiber Structure and Technology

A discussion of the strength of fibers and its structural basis is not possible without a knowledge of fiber structure itself. The essential concepts will be summarized here.

Organic fibers, which account for more than 95% of fiber production, are based on macromolecules or polymers. They may be subdivided into natural fibers such as cotton, wool, etc., with cellulose or proteins as the molecular basis, and fabricated fibers based on either natural or synthetic polymers. In the latter category, polyesters (mainly polyethylene terephthalate), polyamides, polyacrylics, and polyolefins form the few types of polymers used to make the majority of commodity fibers. The production of high-performance specialty fibers (fibers with enhanced thermal or chemical stability, high-strength/high-modulus fibers) makes use of a variety of mainly aromatic or heterocyclic polymers. However, polyethylene can also form a high-strength fiber if processed in a special way (gel spinning).

What sets fibers apart from other polymeric materials is that they are generally highly anisotropic. On the molecular and supramolecular level, this means that there is a high degree of orientation of structural units with respect to the fiber axis. It is worthwhile to consider this in view of the forces acting between the atomic groups making up the material. In macromolecules, covalent bonds having rather high strength are effective along the chain direction, whereas the intermolecular forces perpendicular to the chain direction are relatively weak (van der Waals-type bonds, hydrogen bonding). In order to achieve high strength in the fiber direction, one would like to have the macromolecules arranged extended and parallel to the fiber axis.

Most polymers used for fiber manufacture are partially crystalline, the degree of order in the crystalline regions being much higher than that in the remaining amorphous phase (two-phase model). The alternative microfibrillar concept, where crystalline domains and intrafibrillar disordered domains are coupled in series to form the microfibrils, is used in particular with the natural cellulosic fibers (see *Natural Cellulose Fibers: Properties**). Which model is most appropriate is subject to the nature of the fiber and the property considered.

The spinning process has a decisive influence on the development of a certain morphology. Thermoplastic polymers are usually melt spun, the physical structure of the filaments being controlled by the solidification conditions such as rate of cooling and spinline stress. High spinline stress, achieved by high winding speed, results in increased orientation, and this in turn favors crystallization. Usually the as-spun filaments do not have adequate mechanical properties and a subsequent drawing process slightly above the glass transition temperature is necessary to achieve the required modulus, strength, etc., via further controlled orientation and crystallization. By varying the drawing conditions, fibers with different property profiles can be obtained.

Dry and wet spinning, usually employed with nonthermoplastic polymers (polyacrylonitrile, cellulose, polyaramides), involve dissolution of the starting material in a suitable solvent. After extrusion through an orifice, the solvent is either evaporated (dry spinning) or the extrudate is led through a coagulation bath containing a nonsolvent, where phase separation takes place (wet spinning). In the latter case, a gel fiber is formed, a highly dispersed system of a polymer-rich continuous phase with inclusions of excess liquid. When the solvent is evaporated after drawing, the porous structure collapses, but most likely the original void surfaces remain as weak boundaries separating the material into fibrillar elements. Such morphology is also exhibited by the natural cellulosic fibers (cotton, flax).

A special variant of wet spinning is used for rigid-chain polymers such as p-aramids. In suitable solvents above a certain concentration, nematic solutions are formed. The nematic domains are macroscopically aligned by shear and elongational flow during spinning, and after coagulation fibers having a considerable molecular orientation are obtained. This results in fibers with mechanical properties substantially superior to those of fibers made by melt spinning of flexible polymers.

A completely different way of achieving extraordinarily high chain orientation and extension is by gel spinning—a process successfully realized only with high molar mass polyethylene. The polymer is dissolved in a nonvolatile solvent to form a low-concentration solution. The filaments formed after extrusion (at high temperature) into a cooling bath can be drawn to very high draw ratio, yielding fibers with exceptionally high modulus and strength.

Carbon fibers have a structure similar to the graphite layer lattice. However, their density is appreciably below that of graphite, an indication of the high level of imperfections. Their strength and stiffness depends on the degree of orientation of the layer planes parallel to the fiber axis. The orientation in the perpendicular direction shows a strong local variation. To obtain high-strength/high-modulus carbon fibers, it is imperative to start from precursors that are very well oriented, because this orientation is essentially retained when carbonization and graphitization take place. Glass and ceramic fibers are amorphous and polycrystalline, respectively, with little or no molecular orientation.

2. Mechanical Failure

The usual way to study tensile failure of fibers is by conducting a stress–strain test up to breaking of the fiber (with force and deformation applied along the fiber axis). In most cases this is done at a constant (specified) rate of elongation. A different shape of the stress–strain curve is obtained when experiments are performed at a constant rate of loading, and this should be kept in mind when results of laboratory tests are transferred to technical applications.

Generally load is normalized by the initial cross-sectional area to give stress, the SI units being pascals ($1\,\mathrm{Pa} = 1\,\mathrm{N m^{-2}}$). However, in fiber technology it is more common to normalize by the linear density of the material to obtain specific stress (stress per density). Linear density is expressed in tex ($1\,\mathrm{tex} = 1\,\mathrm{g km^{-1}}$), and specific stress is given in $\mathrm{N\ tex^{-1}}$ ($\mathrm{GPa\,cm^3 g^{-1}}$). Elongation is divided by the length of the undeformed fiber to obtain tensile strain or percentage extension.

Some typical shapes of stress–strain curves of fibers are shown in Fig. 1. The initial slope of the stress–strain curve is the modulus, a material property of the fiber. In contrast, the ultimate properties of elongation at break and strength (stress at break) or tenacity (specific stress at break) are sample properties, showing quite a variation when several samples of the same material are inspected. Whereas material properties are averages over the whole sample volume, ultimate properties reflect the property of the weakest link. As a consequence, the strength of fibers shows a decrease with rising sample length owing to the effect of variability. Some typical data for different fibers are compiled in Table 1.

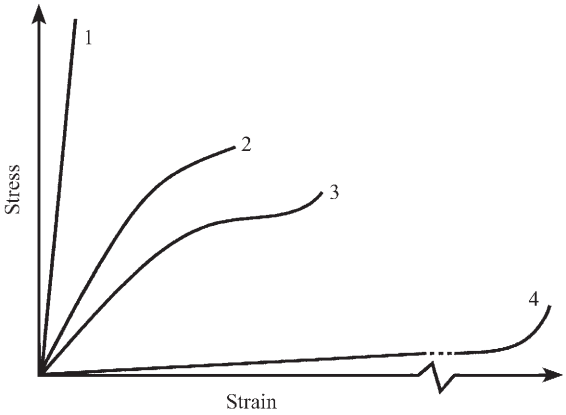


Figure 1
Schematic stress–strain curves of fibers (see text for details).

Table 1
Tensile properties of fibers.

Fiber	Strength (GPa)	Modulus (GPa)
Polyester	0.4–1.3	10–16
Polyamide	0.4–1.0	0.6–6
Polypropylene	0.2–0.6	1–7
Polyacrylonitrile	0.4–0.6	6–9
Rayon	0.3–0.4	8–13
Polyethylene (gel-spun)	2–3	120–170
p-Aramid	2–3	80–140
Carbon	3–4	200–400
Glass	3–5	70–90
Ceramic (Nicalon)	3	200

The different types of stress–strain curves correspond to materials with different types of fracture. Brittle fracture is the typical appearance of failure of glass, carbon, and ceramic fibers. Such fibers have a simple Hookean stress–strain curve (curve 1 of Fig. 1). The break proceeds from a flaw or crack at the fiber surface or (less commonly) in the interior, where peak stresses are generated on loading. When the strength of the material is locally exceeded and the crack starts to grow, the stress at the tip increases even higher giving rise to catastrophic rupture of the fiber. This usually results in a rather flat cleavage plane perpendicular to the fiber axis. There is no indication in the stress–strain curve that the material is approaching the breaking elongation, and the strain is predominantly elastic with no plastic yield.

While brittle fracture is typical for the high-modulus or high-strength reinforcing carbon, glass, or ceramic fibers whose breaking elongation seldom exceeds 1–2%, it is also observed with the highly extensible elastomeric (rubbery) fibers (curve 4 of Fig. 1) having breaking elongations of several hundred percent.

Melt-spun polymeric fibers like polyamide, polyester, and polypropylene show rupture by yield (curves 2 and 3 of Fig. 1). The actual shape of the stress–strain curve reflects how much the fiber is drawn during manufacture. Starting from a flaw, a crack propagates in a controlled manner and opens into a broad notch while the remainder of the fiber is drawn (plastic deformation) under the increasing stress. Eventually, overall failure occurs when the stress on the remaining cross-section exceeds a critical value. Since the plastic deformation is markedly time dependent, the rate of deformation has a significant influence on the appearance of the break.

When the rate of elongation is drastically increased, a different type of failure is observed. This is thought to be owing to a change from isothermal to adiabatic conditions. The heat dissipated on drawing causes a local temperature jump with a softening of the material and rupture by viscous flow. The thermally stimulated molecular relaxation of the polymer gives rise to a mushroom-like broadening of the fracture surface.

Wet-spun fibers such as polyacrylonitrile, rayon, etc., whose stress–strain curves resemble curve 3 of Fig. 1, show breaks with a granular fracture surface. This is a consequence of the coagulation process during fiber formation: The fibrillar elements formed in the gel state on drawing build up fairly independent load-bearing strands. These do not break at exactly the same axial position, but rather at nearby locations in a certain range, thus giving rise to the roughness of the break.

With high-strength wet- or gel-spun fibers like p-aramid and polyethylene, the strength of the fibrillar entities becomes so large that the fracture mechanism turns over to long axial splitting. The weakness of these fibers is in the lateral cohesion of the fibrils, being overcome even by small shear stresses originating at some structural inhomogeneities when the axial tension exceeds a critical level. Up to this level, the stress–strain behavior is practically Hookean (curve 1 of Fig. 1).

Whereas the study of failure through the application of a single excessive load is fairly straightforward and gives reasonable numbers for comparison of different materials, the real situation is often more complicated. Fiber failure can occur after repeated moderate loading over a long period of time. This is termed fatigue. The breaks formed often differ appreciably from that described above, with the detailed conduct of the experiment having a major influence. Owing to the anisotropy of fibers, the direction of the cyclic stresses is important and at least tensile fatigue and flex fatigue need to be distinguished.

See also: Fiber Strength*

Bibliography

Hearle J W S, Lomas B, Cooke W D, Duerden I J 1989 *Fibre Failure and Wear of Materials*. Ellis Horwood, Chichester, UK

Morton W E, Hearle J W S 1993 *Physical Properties of Textile Fibres*. Textile Institute, Manchester, UK

Tomka J G 1989 Textile fibres. In: Allen G (ed.) *Comprehensive Polymer Science*. Pergamon, Oxford, UK, Vol. 2, pp. 487–510

W. Oppermann
Universität Stuttgart, Germany

Thermally Contracting Materials

Thermal contraction in metal (M) oxides is usually the result of the thermal motion of oxygen transverse to M–O–M linkages. However, thermal motion of M transverse to O–M–O linkages can also give rise to thermal contraction. There are only a few examples of this because twofold coordination of metal cations is unusual. However, both Cu_2O and Ag_2O contain linear M–O–M linkages, and both compounds possess negative thermal expansion behavior, as shown by both EXAFS and diffraction studies. Thermal contraction in Cu_2O occurs only below room temperature and is not strong, $\Delta l/l$ $(\alpha) = -2.4 \times 10^{-6}\,K^{-1}$. However, in Ag_2O, thermal contraction is observed from 40 K to 470 K, close to where it decomposes. The thermal contraction is substantial, $\alpha = -6.7 \times 10^{-6}\,K^{-1}$. EXAFS studies show that the Ag–O distances increase and Ag–Ag distances decrease with increasing temperature, consistent with thermal contraction caused by vibration of Ag transverse to the O–Ag–O linkage. Linear O–M–O linkages where M is Cu or Ag also exist in compounds with the delafossite ($CuFeO_2$) structure. The degree of thermal contraction of the O–Cu–O linkage can be substantial, $\alpha = -4.2 \times 10^{-6}\,K^{-1}$ as in $CuScO_2$ below room temperature. In the case of $CuLaO_2$ significant thermal contraction of this linkage continues to the highest measured temperature of 600 K. However, in the case of $AgInO_2$ the thermal contraction of the O–Ag–O linkage is much less than in the case of Ag_2O.

Another, purely electronic, mechanism for thermal contraction is found in some compounds containing Sm or Yb. This mechanism is based on excitation of 4*f* electrons to a conduction band. As more 4*f* electrons are excited with increasing temperature, the rare-earth atoms become smaller, and this can overcome the forces that usually give thermal expansion to yield an overall thermal contraction. The band accepting the 4*f* electrons may be based primarily on 5*d* states of the rare-earth atoms or it can be a band based primarily on states from other atoms in the compound (Fig. 1). For cubic $Sm_{0.75}Y_{0.25}S$ thermal contraction is observed over the entire measured temperature range of 4.2–350 K. In this case the band accepting the 4*f* electrons is the Sm 5*d* band. For hexagonal YbGaGe, the band accepting the 4*f* electrons is the Ga 4*p* band. For the measured temperature range of 100–300 K, thermal contraction occurs along the *a*-axis and thermal expansion occurs along the *c*-axis, giving an overall very small change of volume with temperature. For hexagonal YbCuAl thermal contraction is observed along both the *a*- and *c*-axes below 60 K.

The potential for thermal contraction exists in all oxide framework structures where the oxygen coordination does not exceed two, provided that the framework is not stuffed with cations or molecules. Examples of this are found for tetrahedral MO_2 compounds, $A_2M_3O_{12}$ compounds, cubic AM_2O_7 compounds, compounds with the cubic ZrW_2O_8 structure, and in $AOMO_4$ compounds, which can be monoclinic or orthorhombic. In the case of orthorhombic $TaOVO_4$, thermal contraction occurs along all three crystallographic axes giving an overall thermal expansion coefficient of $\alpha = -5.1 \times 10^{-6}\,K^{-1}$ from 290 K to 900 K. Thermal contraction might be expected in the AO_3 family based on metal oxygen octahedra sharing corners. In fact, most compounds with this structure, such as WO_3 and ReO_3, show only thermal expansion. However, for TaO_2F,

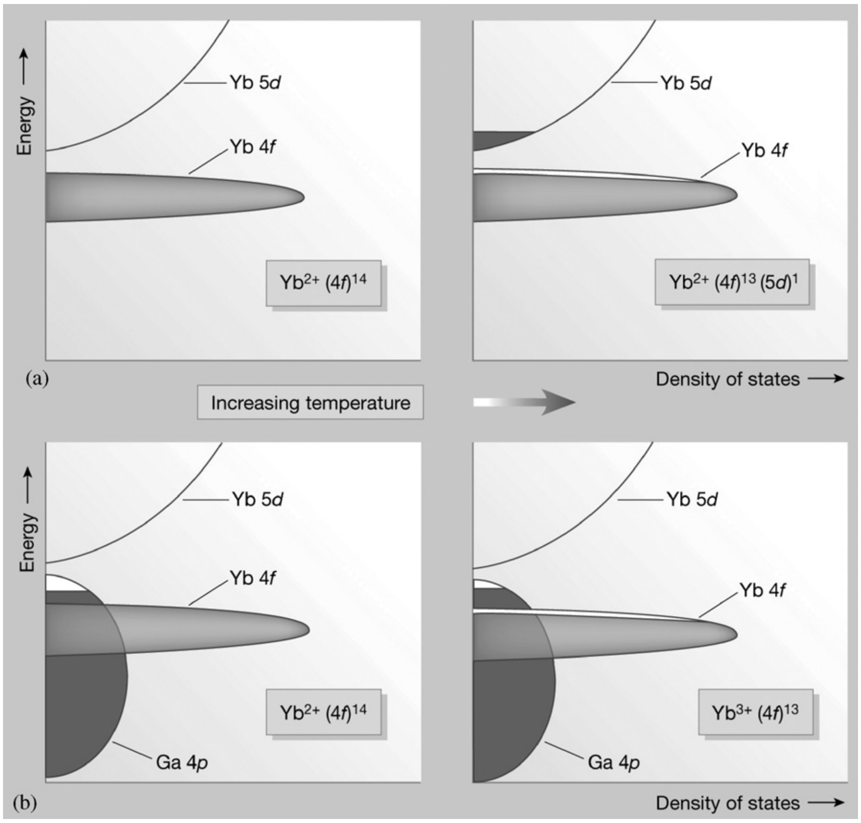


Figure 1
With increasing temperature, electrons are excited from the very narrow Yb (4*f*) band to a broad conduction band. The conduction band may be either the Yb 5*d* band (a) or a band based on another element present in the compound, such as Ga 4*p* in the compound YbGaGe (b). Thermal contraction of Yb occurs in both cases. Reproduced from Sleight (2003) Materials science: Zero-expansion plan, *Nature* **425**, 674–5, with permission.

no detectable expansion or contraction could be detected over the 20–400 K temperature range. This very low thermal expansion is presumed to be the result of nearly exact compensation of the tendencies for thermal contraction and thermal expansion.

Bibliography

Beccara S, Dalba G, Fornasini P, Grisenti R, Sanson A, Rocca F 2002 Local thermal expansion in the cuprite structure: the case of Ag_2O. *Phys. Rev. Lett.* **89**, 255031–4

Li J, Yokochi A, Amos T G, Sleight A W 2002 Strong negative thermal expansion along the O–Cu–O linkage in $CuScO_2$. *Chem. Mater.* **14**, 2602–6

Mattens W C M, Hölscher H, Tuin G J M, Moleman A C, de Boer F R 1980 Thermal expansion and magneto-volume effects in the mixed-valent compound YbCuAl. *J. Magn. Mag. Mater.* **15–18**, 982–4

Mook H A, Holzberg F 1981 Lattice dynamics of the intermediate valence compounds $Sm_{0.75}Y_{0.25}S$ and TmSe. In: Falicov, Hanke (ed.) *Valence Fluctuations in Solids*. Maple, North-Holland, Amsterdam, pp. 113–8

Salador J R, Guo F, Hogan T, Kanatzidis M G 2003 Zero thermal expansion in YbGaGe due to an electronic valence transition. *Nature* **425**, 702–5

Sleight A W 2003 Negative thermal expansion. *Mat. Res. Soc. Proc.* **755**, DD10.6.1–DD10.6.11

Tao J Z, Sleight A W 2003 Very low thermal expansion in TaO_2F. *J. Solid State Chem.* **173**, 45–8

Tiano W, Dapiaggi M, Artioli G 2003 Thermal expansion in cuprite-type structures from 10 K to decomposition temperature: Cu_2O and Ag_2O. *Appl. Crystall.* **36**, 1461–3

A. W. Sleight
Department of Chemistry, Oregon State University, Corvallis, OR 97331-4003, USA

This Page Intentionally Left Blank

W

Whiskers

Filamentary crystals, or whiskers as they are more often known, are usually defined as single crystals having an aspect ratio (ratio of length to diameter) of at least 100. In practice, aspect ratios as high as several thousands are quite common in whiskers.

In the 1930s it was observed that crystals including whiskers often grew from the vapor phase at a supersaturation much too low to allow for growth to proceed by the nucleation of new atomic layers on the growth surface. Frank postulated a crystal growth mechanism that involved a screw dislocation coaxial to the crystal growth axis. This mechanism allowed for very high aspect ratios coupled with growth at low supersaturations, and soon became universally accepted as the mechanism for whisker growth until the advent of the transmission electron microscope (TEM). Since dislocations can be observed with a TEM, workers soon examined the whiskers of several materials looking for the "postulated screw dislocation" at the core or the whisker. Frank's screw dislocation was observed in the whiskers of several materials; however, in the vast majority of instances no dislocation was observed. In those instances in which no dislocation was observed it was simply postulated that two dislocations of unlike sign annihilated each other or that the dislocation responsible for growth glided out of the crystal. In the late 1950s and early 1960s workers reported observing metallic globules attached to the tip of growing whiskers. These "flagpole" whiskers were soon observed to be quite common and in 1964 Wagner and Ellis proposed the vapor–liquid–solid (VSL) mechanism of single-crystal growth. The VLS process soon became the preferred method for the laboratory and commercial growth of whiskers.

A variety of both metal and ceramic whiskers have been growth in the laboratory. Metal whiskers include palladium, silicon, germanium, zinc, cadmium, iron, copper, and silver. Ceramic whiskers grown include aluminum oxide, aluminum nitride, beryllium oxide, molybdenum trioxide, and silicon carbide. Most of the metal whiskers are grown for the purpose of studying the mechanical properties of these fine-diameter single crystals, which are extremely strong with ultimate strengths approaching the theoretical "ideal" limit dictated by atomic cohesion forces. Strengths near the theoretical limit are possible because whiskers have very few structural imperfections and are the closest thing to perfect single crystals achieved. Silicon, germanium, and, at first, silicon carbide whiskers were grown for their semiconducting properties. With the advent of composites, silicon carbide whiskers were also investigated as the reinforcing element in composite structures.

1. *Morphology of Whiskers*

The considerations that dictate the morphology of whiskers are much the same as those for crystals in general. The high-index crystal planes that are less close packed and tend to grow quickly, "grow out" leaving the slower-growing, low-index planes. Both vapor–solid (VS) and VLS whiskers are commonly observed to have either triangular or hexagonal cross-sections. The cross-section is bounded by low-index planes, and the growth direction is most commonly $\langle 111 \rangle$ in cubic systems and $\langle 11\underline{2}0 \rangle$ in hexagonal systems. Growth in the $\langle 111 \rangle$ direction can produce a triangular cross-section or if twinning occurs the cross-section will be hexagonal. Growth in the $\langle 11\underline{2}0 \rangle$ direction produces a hexagonal cross-section.

2. *VS Whiskers*

VS whiskers are so named because growth of the solid whisker occurs by direct condensation from the vapor phase. This growth method is particularly simple for metals with low melting points such as zinc and cadmium. In practice, a temperature gradient is maintained in a gas-tight chamber so as to result in the transfer of material through the vapor phase from the hot end to the cold end. To start the growth process, the chamber, usually quartz, is evacuated and then the hot end containing a crucible of the whisker material is heated so as to maintain a significant vapor pressure. The cold end of the chamber usually contains a seed crystal to encourage nucleation of the whisker and is maintained at a temperature below the melting or sublimation point. For ceramics and metals with higher melting points the process is modified slightly since the necessary vapor pressure must now be achieved through a suitable chemical reaction rather than evaporation. For these systems the crucible at the hot end of the chamber might contain a compound of the whisker material and a reactive gas or a mix of gases is passed through the chamber to produce a volatile species.

The growth of whiskers from the vapor phase was first modeled using the nucleation and growth theories developed in the 1930s. These theories required growth to occur via the nucleation of a "pillbox" that then spreads perpendicular to the growth direction until it covers the entire growth surface (see Fig. 1). Growth therefore required the continuous nucleation of new pillboxes that requires a high thermodynamic driving force or supersaturation. Laboratory experiments soon showed that crystal growth from the vapor occurred at supersaturations much too low to support the continuous nucleation of pillboxes. Frank proposed a mechanism of crystal growth

---

* Cross references marked by an asterisk are included in this volume.

# Cross references marked by a hash can be found by consulting the Encyclopedia of Materials: Science and Technology.

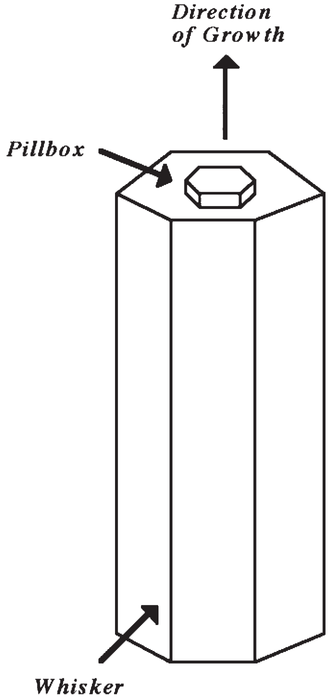


Figure 1
Schematic of whisker growth via nucleation of a pillbox on the tip of a growing whisker. In accord with nucleation and growth theory, whisker growth was thought to occur via the nucleation of pillboxes and growth of the pillbox laterally to cover the tip of the whisker.

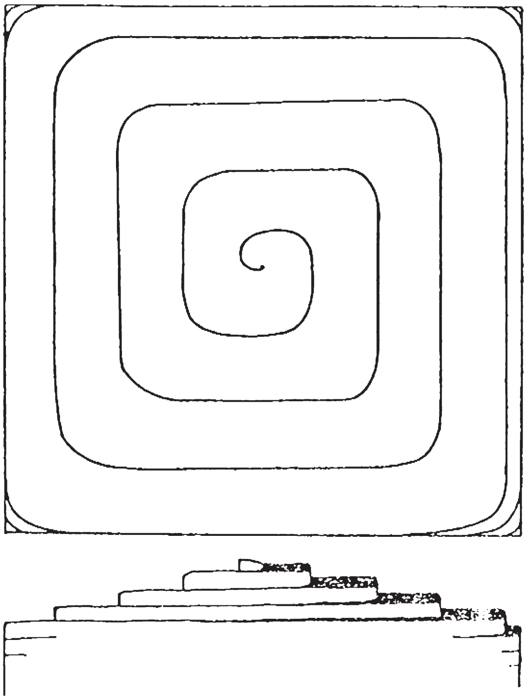


Figure 2
Emergent screw dislocation. A screw dislocation coaxial to the crystal growth axis provides a continuous surface for the attachment of atoms from the vapor. The presence of the screw dislocation allows growth at small driving forces without the need for nucleation (after Burton *et al.* 1951).

based to the (then) new theory of dislocations, which at the time was quite novel given that dislocations had not yet been observed and dislocations were still only "theoretical." In Frank's mechanism a screw dislocation is coaxial with the crystal growth direction. The screw dislocation forms a spiral growth terrace where it exits the growth surface as shown in Fig. 2.

The structural imperfection of the screw dislocation provides a continuous surface for vapor atoms to attach without the need to nucleate "pillboxes" and allows for continuous growth at low supersaturations. If a screw dislocation is present, atom attachment to the growing whisker is easy and the rate-controlling step for whisker growth is generally the rate at which atoms can be supplied from the gas phase.

The screw dislocation central to the Frank mechanism proved difficult to find. One TEM study of whiskers from nine different metals was only able to find a dislocation in whiskers from one of the metals. When dislocations were not observed, which was the vast majority of instances, it was simply postulated that two dislocations of unlike sign annihilated each other or that the dislocation responsible for growth glided out of the crystal. In spite of the paucity of experimental evidence to support the dislocation theory of whisker growth, researchers continued to expound it while ignoring the strange metallic spheres they often observed attached to whisker tips. These metallic spheres would soon be recognized as the hallmark of a new class of whiskers.

3. VLS Whiskers

The discovery of the VLS technique occurred by accident. Liquid metallic spheres were observed to form on the tips of growing whiskers because of the presence of impurities. Impurities tended to concentrate at the whisker tip where they formed low-melting eutectic phases. Wagner and Ellis were the first to identify the VLS process as a separate whisker growth process; however, several previous groups of workers grew whiskers by this method without realizing it. It was soon observed that whiskers with metallic spheres grew faster than those without spheres. The metallic sphere soon became know as a "catalyst" because of its role in accelerating the growth process. The VLS technique is now the preferred technique for growing whiskers and the catalyst is intentionally added to system.

The VLS whisker growth process is shown schematically in Fig. 3. A metallic catalyst is applied to a growth substrate that is heated in a controlled atmosphere furnace. The catalyst melts and beads, and depending on the chemistry, may dissolve material from the substrate until saturation, thus creating a pit. Process gases are then passed and/or generated in the furnace and the catalyst becomes supersaturated and begins to precipitate a whisker. The whisker grows in a preferred orientation and continues growing as long as the supply of process gases continues. Silicon carbide whiskers have been grown up to 10 cm long using this technique. Figure 4 shows a photomicrograph of silicon carbide whisker nucleation.

The metallic catalyst ball at the tip of the VLS whisker (Fig. 5) has a complex function. In addition to being a catalyst for gas-phase reactions, it also serves as a solvent and medium for whisker precipitation. The size of the catalyst ball determines the diameter of the whisker grown, which can vary from less than 1 μm to over 35 μm. For a given catalyst–whisker combination the ratio of the diameter is a function only of the respective surface energies. Whisker growth is quite sensitive to variations in the surface energy of the catalyst ball. Anomalies such as bamboo structures observed in silicon and silicon carbide whiskers are caused by variations in surface energy resulting from perturbations in catalyst chemistry, which in turn is caused by changing gas composition.

The actual size of the catalyst ball varies from the metallic particles applied to the substrate since, in practice, the metallic particles react with the substrate and divide into numerous individual particles of

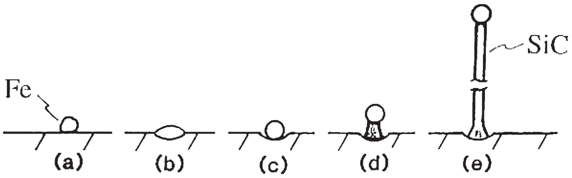


Figure 3
Schematic of VLS whisker growth process. (a) A catalyst (in this case iron) is applied to the substrate. (b) The catalyst particles melt and spread out on the substrate as the temperature of the furnace is raised. (c) The catalyst droplet dissolves substrate material up to the saturation limit and creates a crater. (d) The catalyst absorbs reactants from the gas phase and nucleates a whisker. (e) The whisker achieves the preferred growth direction and grows rapidly.

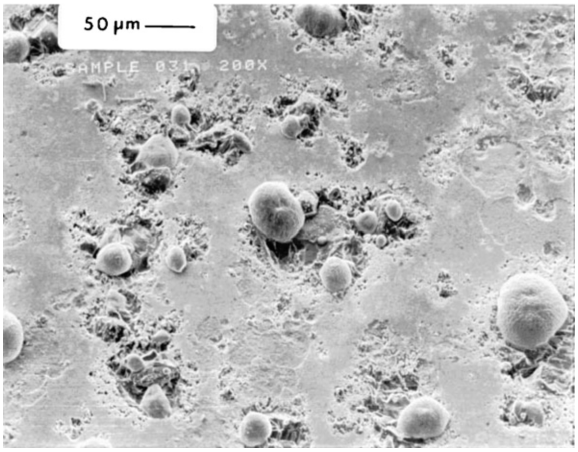


Figure 4
Nucleation of silicon carbide whiskers on a graphite substrate. In this photomicrograph the whiskers have only just nucleated and have not grown to any appreciable extent. The whiskers correspond to stages (c) and (d) in Fig. 3.

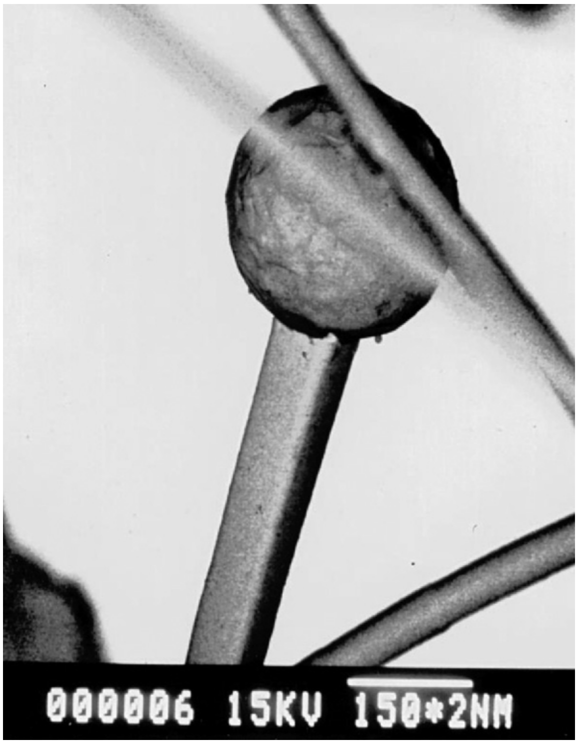


Figure 5
Silicon carbide whisker with catalyst ball. Note triangular profile of whisker (courtesy of J. J. Petrovic).

various sizes giving rise to a range of whisker diameters. The nucleation rate has been shown by Katz to be inversely proportional to the square of the diameter of the catalyst particle; hence, nucleation of submicrometer whiskers occurs rapidly. As the whisker growth process proceeds, however, larger diameter whiskers nucleate and grow, eventually coming to dominate the growth process as the supersaturation in the gas phase decreases and the submicrometer-sized whiskers stop growing. In practice, how fast a whisker grows is determined by the supply of reactants to the catalyst droplet.

Aside from the role impurities played in the discovery of the VLS whisker growth process, they are still essential to the growth of some types of whiskers. For example, nitrogen is added to the gas stream when growing silicon carbide whiskers to promote the formation of the high-temperature β-phase. Nitrogen also gives silicon carbide whiskers their characteristic green color that is evident in Fig. 6. Impurities or the lack of them can cause the color of silicon carbide whiskers to vary from albino (colorless) to blue to black. If surface-active impurities such as sulfur are present they will interfere with the reactions on the catalyst droplet, and if present in large enough quantities shut down the growth process.

Whisker harvesting is usually performed by scraping the growth substrate with a razor blade. In another technique a jet of compressed air is used to fracture the whisker at the base separating it from the substrate. If the whiskers are ceramic and the intended use is in a ceramic composite it is important to remove the metallic catalyst particles. This is accomplished by placing the whiskers in an acid solution to dissolve the catalyst particles.

A commercially important variant of the VLS process involves the production of silicon carbide whiskers from rice hulls. Rice hulls are the waste product left over after the rice is extracted from the kernel. In addition to containing organic matter, rice hulls contain between 15 wt.% and 20 wt.% of silicon. By heating them under carbon monoxide at a temperature of 1350 °C it is possible to produce silicon carbide. Soaking the rice hulls in an iron-containing solution before heating causes VLS whiskers to form. This process was the basis for the first commercially available silicon carbide whiskers.

4. Mechanical Properties of Whiskers

Because of the intense scientific and commercial interest in whiskers their mechanical properties have been the subject of several studies. Table 1 contains

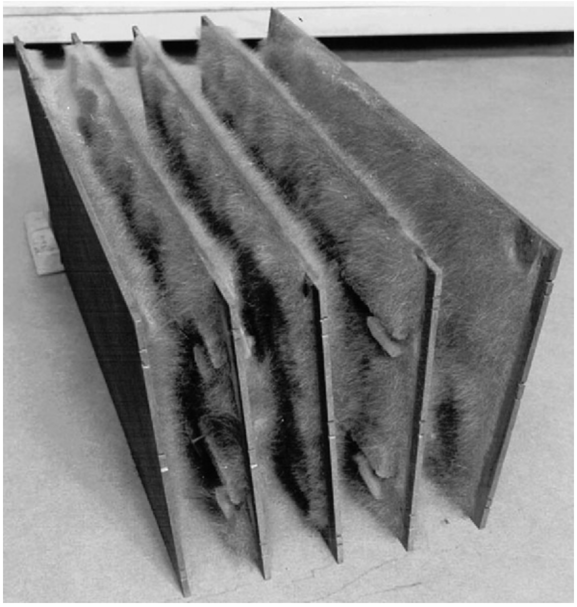


Figure 6
Silicon carbide whiskers on graphite growth substrates awaiting harvesting and further processing.

Table 1
Physical properties of whiskers.

Material	Density (10^3 kg m^{-3})	Melting point (°C)	Tensile strength (GPa)	Young's modulus (GPa)
Aluminum oxide	3.9	2082	14–28	550
Aluminum nitride	3.3	2198	14–21	335
Beryllium oxide	1.8	2549	14–21	700
Boron oxide	2.5	2449	7	450
Graphite	2.25	3593	21	980
Magnesium oxide	3.6	2799	7–14	310
Silicon carbide (α)	3.15	2316	7–35	485
Silicon carbide (β)	3.15	2316	7–35	620
Silicon nitride	3.2	1899	3–11	380

Source: Milewski (1986).

estimated and measured properties for several whiskers. In one study a custom-built microtensile testing apparatus was constructed and used to measure the strength of β-silicon carbide whiskers grown by the VLS process. These whiskers were determined to have tensile strength as high as 23.74 GPa. A theoretical estimate of the tensile strength using the Orowan expression yields a value of 46.3 GPa. The close agreement between the actual measured strength and the theoretical estimate indicates that β-silicon carbide whiskers grown by the VLS process are near-perfect single crystals.

Silicon carbide whiskers have been used as a discontinuous random strengthening element in the fabrication of metal, ceramic, and polymer composites. Silicon carbide whiskers are typically grown to about 2–4 cm in length, and in a few instances to as long as 10 cm. Unfortunately, the typical 2–4 cm length is not long enough to allow their use as a continuous fiber reinforcement element in composites. Whisker felts and tows have been fabricated from 2–4 cm whiskers to allow for their use as a continuous reinforcement element. Coatings have also been applied to whiskers to prevent their reaction with the matrix material and to promote slip on fracture that absorbs additional energy.

5. *Outlook*

The commercial applications and hence the future outlook for whiskers are quite limited. Metallic whiskers were grown only for scientific studies. Semiconducting whiskers such as silicon, germanium, and silicon carbide have not been successfully exploited in electronic devices. A few of the ceramic whiskers such as silicon carbide have found applications as a reinforcing element in ceramic composite materials. For the foreseeable future it seems as though the commercial applications of whiskers will be limited to a few ceramic composite systems.

Bibliography

Bootsma G A, Knippenberg W F, Verspui G 1971 Growth of SiC whiskers in the system SiO_2–C–H_2 nucleated by iron. *J. Cryst. Growth* **11**, 297–309

Burton W K, Cabrera N, Frank F C 1951 *Phil. Trans. R. Soc. London, Ser. A* **243**, 311

Frank F C 1949 The influence of dislocations on crystal growth. *Disc. Faraday Soc.* **5**, 48–54

Givargizov E I 1973 Periodic instability in whisker growth. *J. Cryst. Growth* **20**, 217–26

Lee J G, Cutler I B 1975 Formation of silicon carbide from rice hulls. *Ceram. Bull.* **54**, 195–8

Milewski J V 1986 Whiskers. In: Bever M B (ed.) *Encyclopedia of Materials Science and Engineering*. Pergamon, Oxford, Vol. 7, pp. 5344–6

Petrovic J J, Milewski J V, Rohr D L, Gac F D 1985 Tensile mechanical properties of SiC whiskers. *J. Mater. Sci.* **20**, 1167–77

Wagner R S, Ellis W C 1964 Vapor–liquid–solid mechanism of single crystal growth. *Appl. Phys. Lett.* **4**, 89–90

J. D. Katz
Los Alamos National Laboratory, Los Alamos, New Mexico, USA

Wood Composites: Mineral-bonded

Mineral-bonded composites are composed of materials which combine wood particles or fibers with such minerals as Portland cement, magnesite cements, and gypsum. Examples of products include siding, roof shakes and shingles, tiles, slates, fencing, cladding, tile backer boards, subflooring, and building blocks (Moslemi 1988).

These composites use not only virgin wood fibers and particles, they can also utilize such waste resources as newspapers, magazines, and other solid wastes. In these composites, fibers or particles are held together by the mineral "matrix" generating products that are often water resistant and durable in outdoor applications. Almost all are either fireproof or highly fire-resistant. They combine the characteristics of their component parts, namely wood particles or fibers and the mineral matrix.

1. *Cement-bonded Composites*

Cement-bonded composites utilize Portland cement or other types of cements as the matrix. The proportion of cement in the composite varies from about 20% to over 90% by mass, depending on the type of composite manufactured. Products ranging from flat sheets to building blocks or slabs are made for construction applications.

1.1 Manufacturing Methods

The basic steps of manufacturing process are: preparation of wood or fiber raw materials, mixing these materials with the mineral binder, water and additives, pressing, hardening, or curing (Berkenkamp 1997). Depending on the type of process used, there are substantial differences in the basic layout of the manufacturing plant (Buchmayer 1999). Manufacture of cement-bonded composites is basically divided into three categories: (i) the wet process, (ii) the semidry process, and (iii) the dry process (Kuroki *et al.* 1993).

(*a*) *Wet processes*. These processes rely on a slurry with a great deal of water content to carry and distribute the board-making material. Water is used to develop an aqueous suspension of cement, fibers, and additives. Four different manufacturing techniques are noted here.

The Hatschek Process: This is a continuous operation using a slurry of cement, water, fiber, and additives to form a flat sheet. The forming machine relies on winding very thin sheets (often one millimeter or less in thickness) around an accumulator roll until the target thickness is achieved. At this stage, a knife cuts the "green" sheet, which drops onto a conveyor belt that carries it on for further processing (i.e., pressing, embossing, etc.) (Buchmayer 1999). This is a modification of the process that can also use asbestos fibers and cement in manufacturing a variety of building products.

The Fourdrinier Process: This is similar to the forming concept used in a pulp mill. Both rely on a slurry of cement, water, fibers, and additives, which are poured onto a moving screen conveyor where some of the water is drained. There are two different versions: multilayer or single-layer (Kuroki *et al.* 1993). In the multilayer version the wet sheet is composed of a very thin layer that winds around an accumulator roll in a similar manner to that described for the Hatschek process. In the single-layer design, the mat consists of a single layer, which eliminates the need for an accumulator roll. The single-layer system permits deep embossing but can result in a nonuniform board with more cement precipitating to the bottom of each board formed.

The Batch Process: The fiber–cement slurry is poured into a forming box in a batch system wherein the excess water is drained by pressing the batch. The water exits from the bottom of the batch leaving a malleable cement–fiber mix in the batch. This process can produce deep molds and intricately embossed products.

The Extrusion Process: In this process, the fiber–cement material is extruded through an extrusion machine in a continuous operation. It permits the production of items with complicated cross-sections, hollow-shaped and L-shaped, or corrugated types of product. The process is also in use in connection with asbestos fibers and cement for those regions of the world that permit the use of asbestos.

(*b*) *Semidry processes*. In this type of process, the material used to form the board is relatively dry to the touch. The mixture of cement, wood particles, water, and additives are formed into a mat by machines that use a relatively dry furnish. In fact, these forming machines are very similar to those used for manufacturing resin-bonded particle boards.

Cement-bonded particle board: Wood particles (instead of fibers) are used in combination with Portland cement, water, and additives. Water, cement, and additives are mixed to produce the "furnish" for board manufacture. The furnish is then directed to a forming machine, which deposits the material onto a moving conveyor to produce a mat. The mat is later consolidated by pressing in combination with clamps that hold the pressure for a number of hours to permit the curing of the cement binder. To accelerate the cure, the consolidated mats are placed in elevated room temperatures for some eight or more hours. Upon completion of this step, the clamps are released and boards removed for further processing. This includes curing, drying and sanding. Boards can also be preprimed or finished depending on the intended market. This process permits the use of embossing to produce deep-molded products.

Fast-set cement-bonded board: This process is similar to that of cement-bonded particle board with important exceptions. It does not need long clamp times to permit cement to set in the process described above (Kuroki *et al.* 1993). It accomplishes this by using a mixture of alumina cement and cure accelerators in combination with hot pressing.

Light-weight cement-bonded board: This process reduces board density from about 1250 to approximately $1000\,\mathrm{kg\,m^{-3}}$. It is accomplished by incorporating low-density aggregates into the furnish. For example, the addition of expanded perlite lowers board density. It is important to note that such additions not only lower density but they also affect other board properties (Kuroki *et al.* 1993). In some cases, autoclaving of the boards is incorporated in the manufacturing steps to enhance board properties (strength, dimensional stability, etc.).

(*c*) *The dry process*. Fibers (often asbestos), cement, and additives are dry-mixed. The dry mixture is formed into a mat on a screen conveyer. The loose mat is compressed using a roll press before water is sprinkled onto the mat and, once again, roll-pressed. Cutting of the mat follows this second roll pressing. To produce decorative surfaces, coloring agents or colored silicates are then spread on to the compressed mat followed by a third roll-pressing (which may have a special pattern) followed by curing. Autoclaves are used to enhance product properties by promoting interaction between calcium silicates and calcium hydroxide. Attempts are being made to replace the asbestos fibers with wood fibers. Further research and development, however, is needed before wood fibers can successfully be used in this process.

2. *Magnesia-bonded Composites*

The bonding agent is the basic salt of magnesium chloride, sulfate, nitrate, or phosphate. For composites, magnesia cement, or Sorel cement composed of magnesium salts of hydrochloric, sulfuric, or carbonic acids, are used. Included among magnesia-based binders, for example, are magnesium oxyphosphates, which readily bond lignocellulosic materials in producing composite products (Paszner and Klemarevsky 1989). These cements produce strong bonds when aluminum polyphosphate is reacted with oxides, hydroxides, or carbonates of alkaline metals. The reaction is self-initiating, exothermic, and unaffected by sugars and extractives present in wood which can inhibit Portland cement from hardening (Moslemi 1989). The reaction is rapid. High strength, fire-retardant, and water-resistant composite materials can be manufactured. Products are made by cold-pressing, molding, extrusion, or spraying.

3. Gypsum-bonded Composites

In these composites, gypsum forms the matrix (Lempfer 1995). It is the gypsum hemihydrate ($CaSO_4 \cdot \frac{1}{2}H_2O$) which is obtained by the calcination of natural gypsum, phospho-gypsum (waste gypsum from chemical processes), or FGD (flue-gas desulfurization) gypsum (Lukevich 1997). Basically two different forms of gypsum exist, called alpha- and beta-gypsum. Due to high cost, alpha-gypsum is not used for the manufacture of gypsum-bonded composites. However, beta-gypsum is a low-cost gypsum due to its simple manufacturing process and low energy requirements. Similar to cement-bonded composites, gypsum-bonded composites generally contain a high percentage of gypsum in the product. The gypsum matrix, therefore, participates substantially in the performance of the composite being manufactured. Organic fibers or particles added to gypsum reduce brittleness and improve product toughness.

3.1 Manufacturing Methods

Although there are several methods of manufacturing gypsum fiber and particle boards (Sattler and Lempfer 1989), these processes can be categorized as using either fibers or particles. The former is called gypsum fiber board and the latter particle board. Gypsum fiber boards are made by mixing gypsum with the reinforcing fiber in a relatively dry or slightly moist condition. Not much moisture can be added to the gypsum early in the process due to rapid setting of gypsum. In addition to gypsum and fibers, other materials are often incorporated into the mix (such as gypsum retarders or low-density aggregates) prior to board formation. Once the mixture has been formed into a mat, water is sprayed on this mat just prior to entering a press for mat consolidation. Press time is short due to rapid setting of the gypsum matrix. Once the board exits the press, additional curing and conditioning are performed (Takahashi and Shigekura 1991). During the pressing, embossing can be done through the use of special dies.

The manufacture of gypsum-bonded particle board, with some exceptions,is similar to that described for cement-bonded particle boards. Wood particles are used as the reinforcing element (Sattler and Lempfer 1989).

4. Product Properties

Table 1 summarizes some of the product properties for mineral-bonded composites. Fiber–cement products have emerged as the direct result of attempting to find a replacement for asbestos–cement products. With its adverse impact on human health, the use of asbestos has been discontinued in some industrial nations. In the search for a suitable replacement, a variety of fibers such as steel, glass, synthetic polymers, and cellulosic fibers have been examined. In this search, both technical merits and economic considerations have been evaluated. Cellulosic fibers have emerged as the most technically and economically suitable fiber for fiber–cement products (Ni 1995). In addition to flat sheets, corrugated products are also produced. Gypsum fiberboard is suited not only in applications where traditional wall board is used, but it can be used as a substrate for laminating a variety of surface laminates such as wood veneer, vinyl, foil, etc. It expands market opportunities as compared to traditional wall board. In Europe,

Table 1
Properties of mineral-bonded composites.

Property	Units	Fiber–cement board	Cement-bonded particle board	Magnesia-bonded particle board	Fiber–gypsum board	Gypsum-bonded particle board
Density	kg m^{-3}	1000–1350	1000–1350	1000–1200	900–1000	1000–1200
Bending strength	MPa	6–15	6–15	7–14	4–7	6–9
Modulus of elasticity	GPa	8–16	3–5	3	4–5	3–5
Thickness swelling	% 24h soak	1·0–2·0	1·2–1·8	4–8		
Linear expansion	% 35–90% RH	0·04–0·16	0·12–0·25	0·25	0·04	0·05–0·07

gypsum fiber board in smaller sizes is also used by do-it-yourselfers because of its ease of application (no need for taping, etc.). As Table 1 indicates, gypsum-bonded composites have similar strength properties to those made with cement; however, gypsum is decidedly sensitive to water exposure. An important attribute of most of these composites (made with either cement or gypsum) is their high resistance to fire, an important consideration for construction applications.

5. Blocks and Slabs

These composites combine wood particles with a cementitious binder in the production of building blocks and slabs of relatively low weight (Walter 1999). Some of these products have been produced in Europe since the end of the nineteenth century. Since 1928, wood fiber (shavings and wood wool) products mixed with cement have been on the market under various trade names. Further developments in these technologies have occurred since World War II.

Blocks and slabs are molded and compressed to suit the desired aesthetic and technical requirements. They are made from wood particles which are neutralized and mineralized before being bonded together with cement. Products are relatively lightweight ($750\,\mathrm{kg\,m^{-3}}$), porous, and durable, have high strength-to-weight ratio, good thermal insulation properties, and good sound absorption. They are dimensionally stable and exhibit good freeze/thaw resistance. In addition to blocks and slabs, waffle floor forms, roof panels, acoustic facing panels, and wall components are made.

6. Market Acceptance

Markets for cement-bonded siding and roofing products are expanding in North America and elsewhere. Some of the attributes that contribute to market acceptance are:

* Durability resulting in reduced callback and maintenance
* Resistance to hail and other impacts
* Resistance to high winds
* Lightweight (no need to reinforce roof framing or deck)
* Reduction of waste and scraps
* Resistance to rot, termites, woodpeckers
* Long-lasting, uniform colors
* Dimensional stability especially for fiber–cement and fiber–gypsum products
* Freeze–thaw durability (100-cycle minimum)
* Aesthetic similarity to substitute products
* Code approvals
* Resistance to ultraviolet light
* Comparable installation to all-wood products (saw, nail, etc.)

Given the suitability of mineral-bonded composites for many applications, it is likely that these composites will play a greater role in meeting a variety of construction needs.

For further information the reader is referred to *Proceedings of the Inorganic-bonded Wood and Fiber Composites Conference*. Volumes are issued every two years.

See also: Wood: Nonstructural Panels*; Wood: Nonstructural Panel Processes*

Bibliography

Berkenkamp R 1997 Wood fiber-cement product, process and properties. In: Moslemi A A (ed.) *Inorganic-bonded Wood and Fiber Composite Materials*. Forest Products Society, Madison, WI, pp. 8–13

Buchmayer K 1999 Plant layout and start-up of fiber-cement manufacturing plants. In: Moslemi A A (ed.) *Inorganic-bonded Wood and Fiber Composite Materials*. University of Idaho, Moscow, ID, 6, pp. 99–140

Kuroki Y, Nagatomi W, Yamada J 1993 Manufacture of lightweight cement-bonded particleboard. In: Moslemi A A (ed.) *Inorganic-bonded Wood and Fiber Composite Materials*. Forest Products Society, Madison, WI, 3, pp. 136–42

Kurpiel F 1999 Rapid growth of cement–cellulose fiberboard (CFB). In: Moslemi A A (ed.) *Inorganic-bonded Wood and Fiber Composite Materials*. University of Idaho, Moscow, ID, Vol. 6, pp. 55–77

Lempfer K 1995 Technology trends in gypsum fiberboard. In: Moslemi A A (ed.) *Inorganic-bonded Wood and Fiber Composite Materials*. Forest Products Society, Madison, WI, vol. 4, pp. 41–6

Lukevich L 1997 Flue gas desulfurization gypsum: technology and market development. In: Moslemi A A (ed.) *Inorganic-bonded Wood and Fiber Composite Materials*. Forest Products Society, Madison, WI, pp. 91–5

Moslemi A A 1988 Inorganically-bonded wood composites. *CHEMTECH (August)*, 504–10

Moslemi A A 1989 Wood cement panel products: coming of age. In: Moslemi A A (ed.) *Fiber and Particleboards bonded with Inorganic Binders*. Forest Products Society, Madison, WI, Vol. 1, pp. 12–18

Ni Yong 1995 Natural fiber-reinforced cement composites. Doctoral dissertation, Victoria University of Technology, Australia

Paszner L, Klemarevsky A 1989 New developments in wood bonding with magnesium oxyphosphate cement. In: Moslemi A A (ed.) *Fiber and Particleboards Bonded with Inorganic Binders*. Forest Products Society, Madison, WI, pp. 26–30

Sattler H, Lempfer K 1989 Gypsum-bonded particleboards and fiberboards. In: Moslemi A A (ed.) *Fiber and Particleboards Bonded with Inorganic Binders*. Forest Products Society, Madison, WI, pp. 19–25

Shah S, Peled A, DeFord D, Akkaya Y, Srinivasan R 1999 Extrusion technology for the production of fiber–cement composites. In: Moslemi A A (ed.) *Inorganic-bonded Wood and Fiber Composite Materials*. University of Idaho, Moscow, ID, pp. 61–77

Takahashi K, Shigekura Y 1991 Study on fiberous waste paper-pulp and gypsum. In: Moslemi A A (ed.) *Inorganic-bonded*

Wood and Fiber Composite Materials. Forest Products Society, Madison, WI, pp. 70–7

Walter H 1999 Transforming wood into sustainable materials by the concrete industry. In: Moslemi A A (ed.) *Inorganic-bonded Wood and Fiber Composite Materials*. University of Idaho, Moscow, ID, pp. 349–54

A. A. Moslemi
University of Idaho, Moscow, USA

Wood Products: Decay During Use

Decay of wood during use is a serious problem causing significant losses in strength in stages of decay so early as to be difficult to diagnose even microscopically. The key to preventing decay in wooden buildings is the exclusion of water. Although emphasis in this article is placed on experience with wood in structures, the principles of wood decay are equally applicable to any wood in use.

1. Causal Fungi and their Requirements for Growth

Decay of wood is caused by members of the group of higher fungi called Basidiomycetes related to the common edible mushrooms. These fungi are primitive plants that lack chlorophyll and therefore cannot manufacture their own food, relying instead on the organic matter produced by plants possessing chlorophyll such as the wood produced by trees.

Decay fungi have four basic requirements for growth and production of decay in wood: air, water, a favorable temperature, and a food source. The food is of course the wood itself. This food may be made unavailable to the fungus by poisoning it with a preservative, but normal use of wood in above-ground portions of structures does not usually involve such treatment. Decay fungi have a wider tolerance to temperature than people who occupy wooden structures. Freezing temperatures will stop the progress of decay but will not kill the decay fungi. Temperatures must usually exceed about 60 °C before thermal death occurs. Decay fungi require so little oxygen that it is usually not possible to control decay by limiting this factor in structures; water submersion or deep burial, however, excludes oxygen sufficiently that wood in such conditions is not attacked by decay fungi. Thus, the remaining factor—water—is the key to controlling decay in buildings. Throughout history a major goal of the design and construction of structures has been to shed water effectively. If this is done, and no water leaks are allowed, there is no reason why above-ground portions of wooden buildings should decay. Wood kept dry will never decay.

2. Sources of Inoculum

Any living cells of a fungus, when transferred to a new location favorable for growth, have the potential for reproducing the fungus in this new site. Decay fungi produce spores, which act like seeds of higher plants in that they are capable of starting a new growth of fungus in an area where conditions are favorable. The fungi produce great numbers of such spores light enough to be carried by the winds, thereby assuring that fungi are present wherever conditions conducive to decay exist. Moreover, the careless manner in which wood products often are handled tends to ensure that wood in structures is infected before being placed in service.

Wood may be shipped and stored in the green condition thereby allowing time for it to become infected. One study found that the principal fungi isolated from green lumber were the same ones responsible for most above-ground decay (Dietz and Wilcox 1997, Wilcox and Dietz 1997). Wood products may be delivered to the job site and set on the soil; since decay fungi live in the soil this provides an opportunity for them to move into the wood before it is placed in the structure. If the wood is allowed to dry and remains dry throughout the life of the structure, such preinfection may be of no consequence. However, if conditions later become conducive to decay, the necessary fungi may already be present in a dormant state.

3. Types of Decay

There are two major categories of decay produced by Basidiomycetes: brown rot and white rot. The brown-rot fungi decompose only the carbohydrate fraction of wood, leaving the lignin modified but not metabolized. The residual lignin is darker than cellulose, leading to the brown coloration of the rot, and has very little strength, leading to the excessive shrinkage and cubic checking typical of brown-rotted wood.

White-rot fungi decompose all the major components of wood and can lead to 100% loss in cell wall material. Some of the white-rot fungi attack lignin preferentially in the early stages of decay, while others, sometimes called simultaneous rots, decompose both lignin and carbohydrates at a rate proportional to their original relative concentration in the wood.

4. Microscopic Diagnosis of Decay

Not all decay fungi attack wood in the same way, and a given fungus may produce different effects in different woods. Therefore, not all the features discussed below will necessarily be found in any given piece of decayed wood. However, the following microscopic features may frequently be required for diagnosing early stages of decay; the more they are present, the more confident the diagnosis.

A fungus can be thought of as being like a ball of cotton with filaments running in all directions. The filaments that make up the body of a fungus are called hyphae. Individual hyphae are too small to be seen easily without a microscope, but massed together they can form what is known as a mycelium, which can be seen readily on the wood surface. The hyphae of decay fungi may contain very distinctive loop-like structures called clamp connections. Only Basidiomycetes have them, and the only Basidiomycetes that would be present in wood would be decay fungi. Therefore, if microscopic examination reveals hyphae with clamp connections, then the fungus is a decay fungus. Unfortunately, many decay fungi do not produce clamp connections and some that produce them in culture do not produce them in wood. Clamp connections, then, are a valuable diagnostic tool if present, but their absence supplies no useful diagnostic information.

The hyphae of decay fungi may invade wood in early stages by growth through ray cells and large longitudinal elements, such as resin ducts or vessels. The decay fungi normally leave these elements and ramify throughout the wood. In some cases decay-fungal hyphae may be most numerous in the strength-providing longitudinal elements (fibers or tracheids). Penetration may be through pits or by means of boreholes dissolved through the wood cell walls by the enzymes from the hyphae. The boreholes of decay fungi are usually numerous. They enlarge to several times the hyphal diameter, tend to run in series (serial boreholes) through a number of wood cells walls in a relatively straight line (as seen in radial section) and bear little relation to pits.

Early stages of attack on the cell walls may be characterized by separations within or between cells, and shallow depressions or gouges at the wall–lumen interface. In later stages of decay, white-rotted wood cell walls may appear thinned in a uniform fashion from cell to cell, whereas brown-rotted wood cells appear shrunken and distorted on a very irregular basis from cell to cell. The earliest microscopically detectable stages of brown rot, as low as 1–2% mass loss, are characterized by hyphae in cell lumina, loss of birefringence of secondary walls in cross-sections, serial boreholes, and boreholes smaller in diameter than the hyphae (Wilcox 1993a 1993b).

White-rot fungi decompose all the major cell-wall components and some remove them at rates approximately proportional to the amounts present in sound wood, thereby leading to the observed progressive thinning of the cell walls.

Brown-rot fungi decompose primarily the carbohydrate fraction of the wood. Lignin is altered by their attack, but little of it is removed. In early stages of brown rot the wood constituents may be depolymerized more rapidly than the decomposition products are metabolized by the fungus, leading to a buildup of low-molecular-weight molecules which constitute mass but provide no strength.

Table 1
Expected strength loss at an early stage of decay (5–10% mass loss) in brown-rotted softwoods. Because of the drastic reduction in most strength properties caused by decay at such an early stage as to be difficult to diagnose, it is easy to see that it is considerably safer to prevent decay in the first place than to rely upon ability to detect and diagnose it accurately in sufficient time to effect a remedy before failure occurs.

Strength property	Expected strength loss (% of strength of sound wood)
Toughness	$\geqslant 80$
Impact bending	80
Static bending (MOR and MOE)	70
Compression perpendicular to grain	60
Tension parallel to grain	60
Compression parallel to grain	45
Shear	20
Hardness	20

Source: Wilcox (1978).

5. *Effects of Decay on Strength Properties*

Resistance to impact loading is the mechanical property of wood most sensitive to the presence of decay. Toughness, for example, may be largely destroyed at a stage of decay so early that it may be hard to detect even microscopically (Table 1). Those strength properties relied upon in the use of wood as a beam, a common mode of service, are also severely diminished in very early stages of decay. In such early decay stages there is little difference noted between brown rot and white rot with regard to the magnitude of strength losses. However, as decay advances, further effects on strength are much more severe in brown rot than in white rot.

See also: Wood, Constituents of*; Wood: Preservative Treated#; Wood Products: Chemical Degradation#; Wood Products: Deterioration by Insects and Marine Organisms#; Wood: Strength and Stiffness*.

Bibliography

Dietz M, Wilcox W W 1997 The role of pre-infection of green douglas-fir lumber in aboveground decay in structures in California. *For. Prod. J.* **47** (5), 56–60

Eslyn W E, Clark J W 1979 *Wood Bridges: Decay Inspection and Control*. USDA Agriculture Handbook No. 557. US Government Printing Office, Washington, DC

Graham R D, Helsing G G 1979 *Wood Pole Maintenance Manual: Inspection and Supplemental Treatment of Douglas-Fir and Western Redcedar Poles*. Research Bulletin 24. Oregon State University, Corvallis, OR

Meyer R W, Kellogg R M (eds.) 1982 *Structural Uses of Wood in Adverse Environments*. Van Nostrand Reinhold, New York

Nicholas D D (ed.) 1973 *Wood Deterioration and its Prevention by Preservative Treatments*. Syracuse University Press, New York

Rosenberg A F, Wilcox W W 1982 How to keep your award-winning building from rotting. *Wood Fiber* **14**, 70–84

Scheffer T C, Verrall A F 1973 *Principles for Protecting Wood Buildings from Decay*, USDA Forest Service Research Paper FPL 190. US Forest Products Laboratory, Madison, WI

US Department of Agriculture 1987 *Wood Handbook: Wood as an Engineering Material*, USDA Agriculture Handbook No. 72. US Government Printing Office, Washington, DC

Wilcox W W 1978 Review of literature on the effects of early stages of decay on wood strength. *Wood Fiber* **9**, 252–7

Wilcox W W 1993(a) Comparative morphology of early stages of brown-rot wood decay. *IAWA J.* **14**(2), 127–38

Wilcox W W 1993(b) Comparison of scanning electron microscopy and light microscopy for the diagnosis of early stages of brown rot wood decay. *IAWA J.* **14**(3), 219–26

Wilcox W W, Dietz M 1997 Fungi causing above-ground wood decay in structures in California. *Wood Fiber Sci.* **29**(3), 291–8

Zabel R A, Morrell J J 1992 *Wood Microbiology: Decay and its Prevention*. Academic Press, San Diego, CA

W. W. Wilcox
University of California, Berkeley, California, USA

Wood Products: Weathering

The surface degradation that occurs when wood is used outdoors and above ground is termed weathering. Weathering should not be confused with decay caused by basidiomycete fungi (see *Wood Products: Decay During Use**), which can significantly reduce the strength of structural timber (Feist 1990). Because weathering is a superficial phenomenon its effects on the mechanical properties of wood are small and, accordingly, there are examples of wooden buildings such as the stave (pole) churches in Norway that are still structurally sound despite having been exposed to weathering for over 900 years (Borgin 1970). The most obvious features of weathered wood are its gray coloration and rough surface texture (Fig. 1). Microorganisms colonize weathered wood, but conditions at exposed wood surfaces generally do not favor decay. Hence, the defining features of weathering are its superficial nature and the minor role of microorganisms compared to environmental factors in degradative processes. This article reviews the weathering of wood, with emphasis on the causal agents of weathering, the effects of weathering on wood properties and performance, and, finally, the protection of wood from weathering. The first scientific article on the weathering of wood appeared in the nineteenth century (Berzelius 1827 cited in Feist and Hon 1984), and extensive reviews have been published by Kalnins (1966), Feist and Hon (1984), and Feist (1990).

1. Mechanisms and Environment

The main environmental factors that cause weathering are: solar radiation (ultraviolet and visible light), oxygen, water, heat, particulate matter, and environmental pollutants. The maximum amount of solar radiation available at the earth's surface on a clear day is normally 1000 $\mathrm{W m^{-2}}$. The composition of such radiation is approximately 5% ultraviolet (286–380 nm), 45% visible (380–780 nm), and 50% infrared (780–3000 nm). In order to act upon wood, solar radiation must be absorbed by one of wood's chemical constituents. Experimentation has shown that the aromatic lignin component of wood (see *Wood, Constituents of**) strongly absorbs ultraviolet (UV) light with a distinct maximum at 280 nm and a tail extending beyond 400 nm into the visible portion of the spectrum (Kalnins 1966). The heartwood of many wood species also absorbs light beyond 500 nm because of the presence of low molecular weight extractives such as flavonoids, tannins, stilbenes, and quinones. In the case of synthetic homopolymers, photodegradation can be represented in a simplified manner as follows. Energy from absorbed radiation (*) can be dissipated in the polymer through the cleavage of molecular bonds (photolysis) resulting in the formation of a free radical ($R^* \rightarrow R\bullet$); that is a molecule which is highly reactive because it has an unpaired valence electron. Once a free radical has formed it can readily react with atmospheric oxygen to form a peroxy radical ($R\bullet + O_2 \rightarrow ROO\bullet$). The peroxy radical is capable of attacking the polymer backbone (RH) via hydrogen abstraction, forming a hydroperoxide and another free radical ($ROO\bullet + RH \rightarrow ROOH + R\bullet$). The hydroperoxide is very unstable to UV radiation and undergoes photolysis forming additional free radicals ($ROOH \rightarrow RO\bullet + \bullet OH$). The increasing accumulation of free radicals results in extensive cleavage of molecular bonds and consequent loss of the physical properties of the polymer. The photodegradation of wood proceeds in an analogous manner except that its degradation is more complicated because it consists of a blend of polymers (lignin, cellulose, and hemicellulose) and low molecular weight extractives that differ in their susceptibility to solar radiation. Furthermore, the precise mechanisms and reaction pathways involved in the photodegradation of each of these components have yet to be elucidated. However, it is clear that the key step involved in the photodegradation of wood is

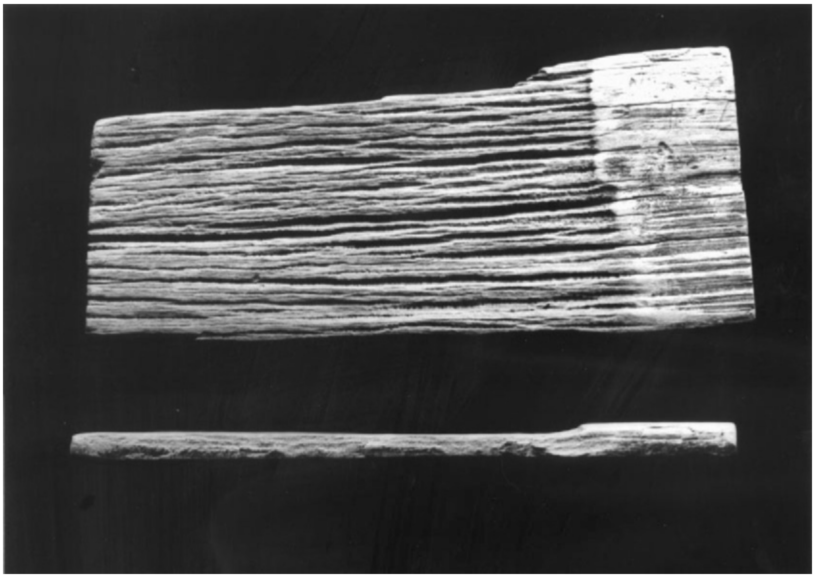


Figure 1
Weathered Norfolk Island pine (*Araucaria heterophylla* (Salisb.) Franco) roofing shingle, measuring 160 × 70 mm, from a nineteenth century building in Norfolk Island, Australia: note the severe checking (top) and erosion (bottom) of the exposed surface.

photolysis and fragmentation of lignin resulting in the formation of free radicals (Feist and Hon 1984). These free radicals may then cause further degradation of lignin and photooxidation of cellulose and hemicellulose. Free radical reactions may be terminated by reaction of radicals with photodegraded lignin fragments forming colored unsaturated carbonyl compounds, which explains why wood yellows when exposed to light. In accord with energetic considerations, the UV portion of the solar spectrum is most effective in causing degradation of wood, but visible light is also involved in weathering. Thus Derbyshire and Miller (1981) noted that "wood exposed only to wavelengths greater than 400 nm will degrade at about half the rate of material exposed to the full solar spectrum."

Water has an important role in the weathering of wood. The leaching of photodegraded lignin and hemicellulose fragments from weathered wood surfaces by rain is in part responsible for the characteristic gray color of weathered wood. Dimensional change caused by the wetting and drying of wood generates surface stresses that cause checking and warping of timber (Feist 1990). The photodegradation of wood also proceeds more rapidly in the presence of moisture, possibly because water molecules swell wood thereby opening up inaccessible regions of the cell wall to degradation (Feist and Hon 1984).

Heat accelerates the chemical reactions involved in the weathering of wood. Exposure of wood to low temperatures and repeated freezing and thawing can also contribute to the physical deterioration of wood. An additional factor involved in the weathering of wood in cold climates is abrasion by windblown particles of ice. For example, the Australian Antarctic explorer Mawson (1915 pp. 123–4) wrote: "the abrasion-effects produced by the impact of the snow particles were astonishing … A deal (*Pinus* sp.) box, facing the wind, lost all its painted bands and in a fortnight was handsomely marked; the hard knotty fibres being only slightly attacked, whilst the softer, pithy laminae were corroded to a depth of one-eighth of an inch [approximately 3 mm]." Windblown sand and salt can also cause similar, if less spectacular, abrasion of wood (Feist 1990).

The main pollutants in the atmosphere are dust and smoke particles and volatile pollutants including sulfur compounds, ammonia, nitrogen oxides, carbon monoxide, and saturated and unsaturated aliphatic and aromatic hydrocarbons and their derivatives. There have been relatively few studies of the effects of these pollutants on the weathering of wood, but observations in the field and *in vitro* experiments have all tended to suggest that the weathering of wood is more rapid in polluted than in unpolluted atmospheres (Williams 1987).

The frequent wetting and drying and high temperatures and levels of solar radiation at exposed wood surfaces are unfavorable to microbial activity. Nevertheless bacteria and certain microfungi are frequently observed colonizing weathered wood, where they contribute to its gray coloration. The dimorphic black yeast *Aureobasidium pullulans* (de Bary) Arnaud in particular is frequently isolated from weathered wood. *A. pullulans* is capable of withstanding temperatures of 80 °C, growing over a pH range of 1.9 to 10.1 and surviving for long periods without moisture. Hence it is particularly well suited to the microenvironment of weathered wood (Schmidt and French 1976). *A. pullulans* metabolizes photodegraded lignocellulose, but organisms capable of enzymatically degrading wood also colonize weathered wood. However, their ability to cause significant degradation is, as mentioned in the introduction, limited by conditions at weathered wood surfaces.

2. *Effects of Weathering*

Wood exposed to the weather changes color very rapidly. Light colored woods, including most coniferous species, darken in color and become yellow or brown due to the accumulation of photodegraded lignin compounds in the wood. Dark colored woods that are rich in phenolic extractives may fade initially before becoming yellow or brown. Irrespective of these initial color changes, wood exposed outdoors for 6–12 months (depending on climatic conditions) becomes gray as photodegraded lignin fragments are leached from the wood resulting in surface layers that are rich in cellulose. Wood exposed outdoors in coastal or very dry environments is often silvery-gray in color, but in other environments weathered wood has a dark gray, blotchy appearance due to the presence of fungal spores, hyphae, and pigments within surface wood layers.

The rapid erosion of low-density earlywood often gives weathered wood a corrugated appearance (Fig. 1). The rate of erosion of wood during exterior exposure largely depends on its density, and accelerated weathering studies have shown a linear relationship, within the density range 0.3–1.0 g cm^{-3}, between wood density and erosion (Sell and Feist 1986). Accordingly, low density species such as western red cedar (*Thuja plicata* D. Don) erode at a rate of 12 mm per century when exposed vertically facing south in the northern hemisphere, whereas comparable figures for higher density softwoods such as Douglas fir (*Pseudotsuga menziesii* (Mirb.) Franco.) and high density hardwoods are 6 mm and 3 mm, respectively (Feist 1990). In addition to erosion, the surface texture of weathered wood is further degraded by the formation of checks. Photodegradation of wood and stresses generated by wetting and drying result in the formation of macroscopic checks where adjacent cells or tissues differ in cell wall thickness or strength. For example, checks often develop at growth ring boundaries and at the interfaces between rays and tracheids. Microscopic changes to the structure of wood precede any evidence of macroscopic damage during weathering. The lignin-rich middle lamella that bonds adjacent tracheids and fibers together is rapidly eroded during weathering, and adjacent primary and secondary cell wall layers show progressive thinning with increasing exposure (Fig. 2(a)–(d)). The most obvious changes to the microscopic structure of longitudinal surfaces are the formation of microchecks originating in bordered and half-bordered pits (Fig. 2(e)–(f)) and the degradation of ray tissues.

All of the major chemical constituents of wood are degraded during weathering. Lignin is depolymerized and, as mentioned above, low molecular weight lignin fragments are leached from wood by rain. The use of spectroscopic techniques, such as Fourier transform infrared spectroscopy (FTIR) and x-ray photoelectron spectroscopy, that can probe the chemical composition of surfaces, has shown that the degradation of lignin at exposed wood surfaces is extremely rapid. For example, FTIR spectroscopy of weathered radiata pine (*Pinus radiata* D. Don) veneers showed a remarkably rapid decrease in the peak at 1505 cm^{-1}, which corresponds to aromatic C=C bond stretching in lignin. Spectra suggested perceptible surface (1–2 μm) delignification after 4 hours exposure, substantial delignification after 3 days, and almost complete surface delignification after 6 days (Fig. 3) (Evans *et al.* 1996). Hemicelluloses, particularly those containing xylose and arabinose, are also degraded and leached from weathered wood, and hence, as mentioned above, weathered wood surfaces are rich in cellulose. Because of this it was assumed for many years that cellulose was less affected by weathering than the other chemical constituents of wood, but viscometry studies of cellulose isolated from weathered wood have shown that cellulose in wood is rapidly depolymerized when exposed to the weather (Evans *et al.* 1996).

The effects of weathering on the structure and chemical composition of wood are superficial in nature. Hence, processing technologies and applications of wood that depend on its surface properties are severely affected by weathering. Notable in this regard is the painting and finishing of wood. Weathering of wood for 2–4 weeks prior to painting has been shown to significantly reduce the adhesion and performance of finishes applied to wood (Williams *et al.* 1987). Weathering also reduces the natural durability of western red cedar roofing shakes by leaching fungitoxic thujaplicins from wood. Because weathering is confined to wood surface layers the mechanical properties of wood, assuming decay to be absent, are largely unaffected by prolonged exposure to the weather. In contrast the mechanical properties of wood composites, which depend in part on the strength of wood–adhesive bonds, can be significantly reduced by moisture-induced dimensional changes when they are used outdoors. Despite these deleterious effects of weathering on the utility of wood, in certain specialized applications weathered wood is preferred to fresh wood. A good example of this is the use of weatherboards for the construction

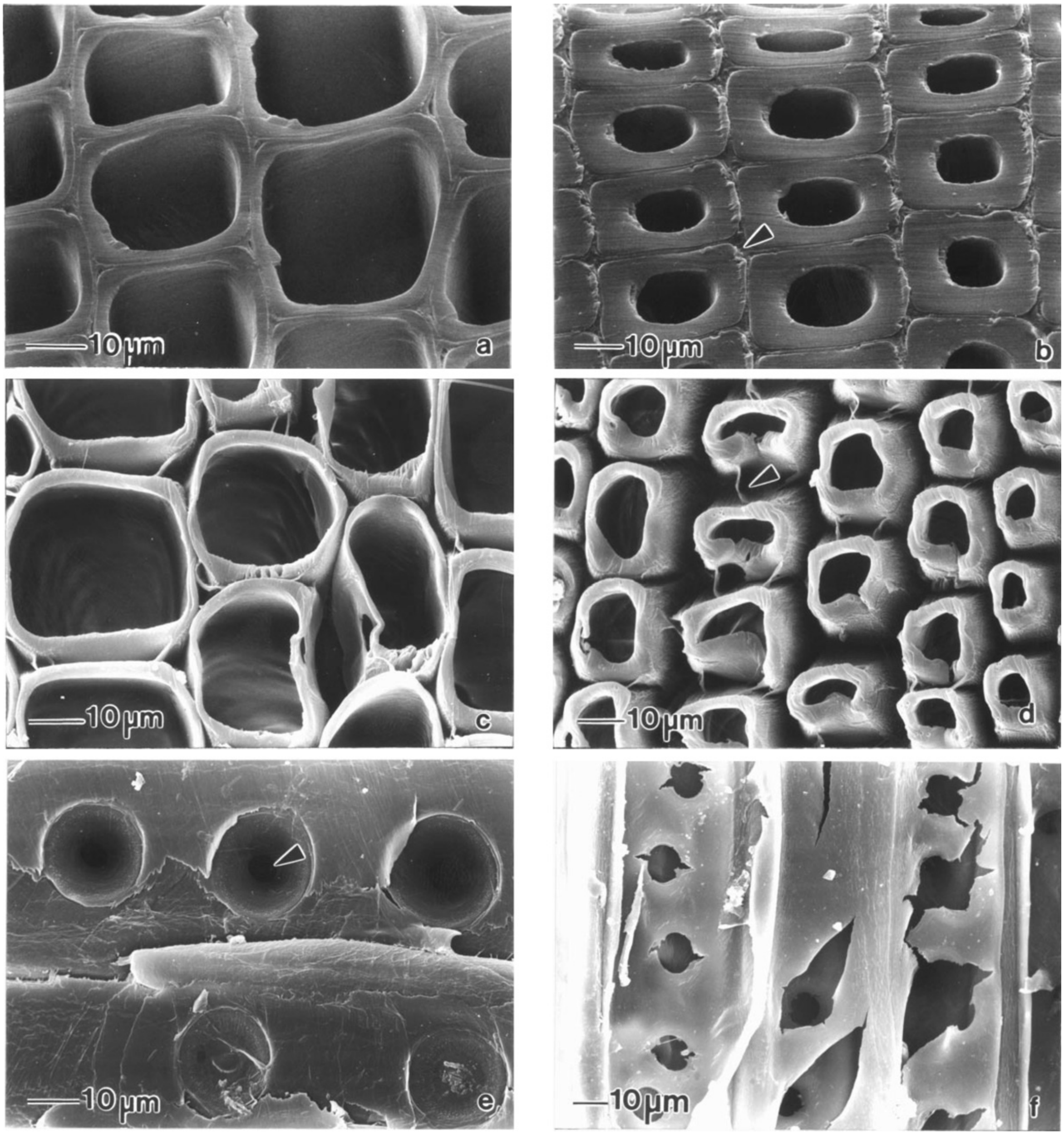


Figure 2
Effects of weathering on the microscopic structure of Scots pine (*Pinus sylvestris* L.) wood. (a) Unexposed earlywood tracheids with thin undamaged cell walls and large cell lumens. (b) Unexposed latewood tracheids with thick cell walls and small lumens (note the middle lamella that cements tracheids together, arrowed). (c) Earlywood tracheids exposed to the weather for 30 days (note erosion of middle lamellae, particularly at cell corners and thinning of cell walls). (d) Latewood tracheids exposed to the weather for 30 days (note erosion of middle lamellae and delamination of cell walls, arrowed). (e) Unexposed earlywood tracheids showing the apertures of bordered pits that allow water in conifer trees to flow from one cell to another, arrowed. (f) Earlywood tracheids exposed to the weather for 30 days (note microchecking originating in bordered pit apertures).

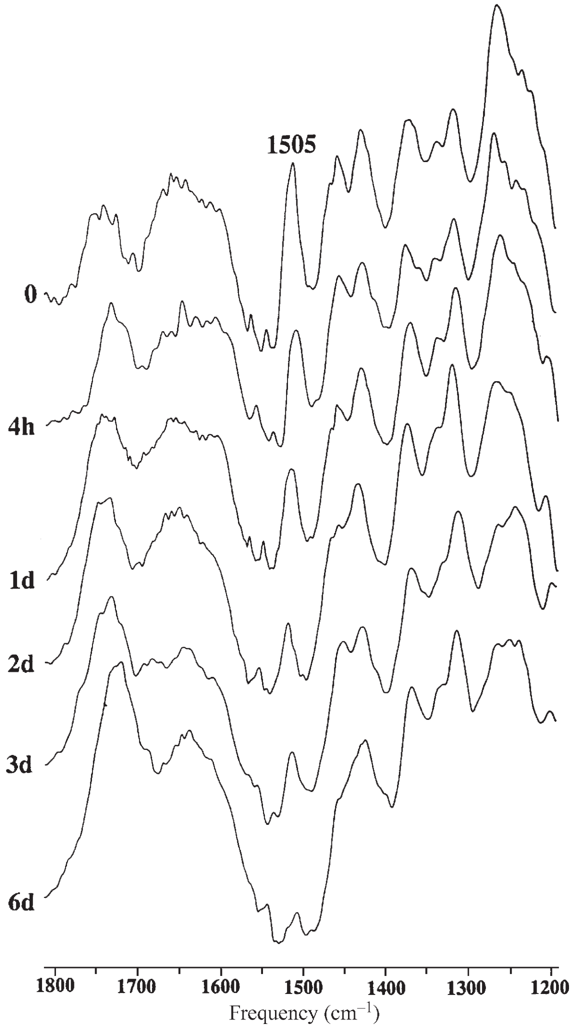


Figure 3
Fourier transform infrared internal reflectance spectra of radiata pine veneers exposed to the weather for periods ranging from four hours (4 h) to six days (6 d) compared to the spectrum for an unexposed control (0).

of "New England type" barns, where the wood may be artificially weathered prior to building construction in order to give the building an aged appearance in keeping with its rural surroundings (Anon 1976).

3. Protection

The most common method of protecting wood from weathering is through the use of coatings (see *Wood: Finishes and Coatings*$^{\#}$). Paints contain pigments that screen wood from solar radiation and because they form a film over the wood surface they prevent surface wetting and erosion. A correctly applied and maintained paint system including a primer and at least two topcoats can greatly reduce the deleterious effects of weathering on wood. However, paints are less effective in controlling decay and dimensional movement and therefore they often perform better on wood that has been pretreated with a water-repellent preservative. The high maintenance requirements of paints and their tendency when used on nondurable timber to trap water and encourage decay has led to increased use of penetrating water repellent stains as a means of protecting wood used outdoors. Stains contain a variety of additives to reduce the weathering of wood, including pigments and UV stabilizers to screen wood from solar radiation, hydrophobic additives to repel water and reduce checking, and a biocide to control the growth of microorganisms. Stains provide protection against weathering for 3–6 years depending on wood species and surface texture, type and quantity of stain applied to the wood, and degree of exposure to the weather. Paints and to a lesser extent stains modify the appearance of wood. For end uses where it is important to retain the wood's natural color or texture the wood can be finished with a clear coating. Clear coatings, however, although they often contain UV stabilizers and a biocide, are limited in their ability to protect wood from weathering because they transmit light, which can degrade the underlying wood surface. Hence they perform badly on wood used outdoors and invariably fail by peeling and cracking within 2–3 years of application. One means of increasing the performance of clear finishes on wood is to photostabilize the underlying wood surface prior to application of the clear finish. Pretreatment of wood with dilute aqueous solutions of transition metal compounds (chromium, manganese, titanium, copper), and particularly chromium trioxide, can photostabilize wood and increase the longevity of clear coatings applied subsequently to the wood. Health concerns about the use of hexavalent chromium, however, have discouraged commercial development of this concept except in Japan where pretreatment of wooden doors with chromium trioxide was used in the 1980s to enhance the weathering resistance of acrylic-urethane finishes. When wood is treated with chromium trioxide, lignin phenolic subunits at wood surfaces are modified producing oxidized lignin complexes with considerable light, thermal, and solvolytic stability (Schmalzl *et al.* 1995). It has been suggested that such complexes could function as an "ultraviolet screen" protecting wood from the effects of the weather. Chemical treatments or additives such as UV absorbers and hindered amine light stabilizers that achieve similar photoprotective effects may also be useful in improving the performance of clear finishes. Chemical bonding or grafting of UV absorbers to wood is particularly effective in photostabilizing wood and improving the performance of clear finishes (Kiguchi and Evans 1998), but commercialization of such treatments awaits the development of less costly UV absorbers that can be more easily bonded to wood. Until more effective methods of photostabilizing wood have been developed it is unlikely that clear finishes on wood will be able to match the exterior performance of paints and other pigmented finishes.

4. *Future Directions*

Architects and engineers are increasingly insisting on materials that offer longer service life and lower repair and maintenance costs. In addition there is pressure to use materials that can be recycled and produced on a sustainable basis. Wood is a renewable material that can be easily repaired and in most cases recycled. However, weathering reduces the service life of wood and increases its maintenance costs. For example, premature failure and replacement of chromated copper arsenate-treated pine poles and decking are often caused by checking, and the higher maintenance costs of exterior wooden joinery compared to substitutes made from unplasticized polyvinyl chloride or aluminum are caused in part by weathering-induced failure of surface coatings. Hence there is likely to be continuing interest in the development of additives and treatments to reduce the weathering of wood. A deeper understanding of the mechanisms involved in photodegradation of wood could lead to more economical and effective methods of photostabilizing wood, which could significantly improve the durability of clear finishes for timber. Wood is increasingly being converted into a wide range of composites and being combined with a diverse range of other materials including thermoplastic and thermosetting polymers, ceramics, and other lignocellulosics. The evaluation of the weathering resistance of such materials and the development of appropriate treatments to enhance their durability will become increasingly important in future.

Bibliography

Anon 1976 New wood with an old look. *Wood Wood Prod.* **81**, 33–5

Borgin K 1970 The stability, durability, and weather resistance of wooden houses under cold climatic conditions. *Meded. Fak. Bosb. Univ. Stellenbosch.* **4**, 6

Derbyshire H, Miller E R 1981 The photodegradation of wood during solar irradiation. Part 1. Effects on the structural integrity of thin wood strips. *Holz Roh-Werkst* **39**, 341–50

Evans P D, Thay P D, Schmalzl K J 1996 Degradation of wood surfaces during natural weathering. Effects on lignin and cellulose and on the adhesion of acrylic latex primers. *Wood Sci. Technol.* **30**, 411–22

Feist W C 1990 Outdoor wood weathering and protection. In: Rowell R M, Barbour R J (eds.) *Archaeological Wood: Properties, Chemistry, and Preservation.* American Chemical Society, Washington, DC, Chap. 11, pp. 263–98

Feist W C, Hon D N-S 1984 Chemistry of weathering and protection In: Rowell R M (ed.) *Chemistry of Solid Wood.* American Chemical Society, Washington, DC, Chap. 11, pp. 401–51

Kalnins M A 1966 Surface characteristics of wood as they affect the durability of finishes. Part 2, Photochemical degradation of wood. *US For. Prod. Lab. Report* **57**, 23–60

Kiguchi M, Evans P D 1998 Photostabilisation of wood surfaces using a grafted benzophenone UV absorber. *Polym. Degrad. Stab.* **61**, 33–45

Mawson D 1915 *The Home of the Blizzard. Being the Story of the Australasian Antarctic Expedition, 1911–1914.* Heinemann, London

Schmalzl K J, Forsyth C M, Evans P D 1995 The reaction of guaiacol with iron III and chromium VI compounds as a model for wood surface modification. *Wood Sci. Technol.* **29**, 307–19

Schmidt E L, French D W 1976 *Aureobasidium pullulans* on wood shingles. *For. Prod. J.* **26**, 34–7

Sell J, Feist W C 1986 Role of density in the erosion of wood during weathering. *For. Prod. J.* **36**, 57–60

Williams R S 1987 Acid effects on accelerated wood weathering. *For. Prod. J.* **37**, 37–8

Williams R S, Winandy J E, Feist W C 1987 Paint adhesion to weathered wood. *J. Coat. Technol.* **59**, 43–9

P. D. Evans
Australian National University, Canberra, Australian Capital Territory, Australia

---

Wood, Constituents of

Wood is a cellular lignocellulosic material consisting of cellulose embedded in a hemicellulose/lignin matrix. The composition of wood varies at all levels from species to species, among cell types, and within the cell wall itself. The chemical composition of softwoods (conifers) differs from that of hardwoods (arboreal angiosperms) in the structure and content of lignin and hemicelluloses (Table 1). Wood may also contain significant quantities of extractives (resins and gums) in the form of low molecular weight extracellular compounds. The inorganic component of wood is typically less than 0.5%.

1. *Cellulose*

Cellulose is one of the most abundant organic compounds on earth. It is a linear homopolysaccharide made up of β-1,4-linked D-glucopyranoside residues in the $^4C_1$ chair conformation (Fig. 1, (**1**)). In both softwoods and hardwoods the cellulose content is approximately 42%. In the native state, adjacent parallel cellulose chains are aggregated together, by inter- and intramolecular hydrogen bonds, to form microfibrils approximately 4nm wide. These microfibrils contain both highly ordered (crystalline) and less

Table 1
Chemical composition of normal and reaction woods in typical softwood and hardwood species.

	Radiata pine (*Pinus radiata*)		Birch (*Betula papyrifera*)	
Component	Normal wood (%)	Compression wood (%)	Normal wood (%)	Tension wood (%)
Cellulose	40	36	42	50
O-Acetylgalactoglucomannan	20	12		
Glucomannan			3	2
1,4-Galactan		8		8
Arabino-4-*O*-methylglucoronxylan	11	11		
O-Acetyl-4-*O*-methylglucoronxylan			30	22
Lignin	27	32	24	17
Extractives	2	1	1	1

ordered (amorphous) regions. The native crystalline structure of cellulose is classified as the cellulose *I* allomorph. The measured degree of polymerization of native cellulose is of the order of 10 000. The microfibrils are further aggregated with lignin and hemicelluloses to form macrofibrils. The high level of bonding in cellulose micro- and macrofibrils results in both high tensile strength and general insolubility in most solvents.

2. *Hemicelluloses*

Classification of plant polysaccharides has been based on their extraction methods. Hemicellulose is a term used for plant cell wall polysaccharides, other than cellulose and pectin, which are soluble in alkali solutions. Hemicellulose consists of various heteropolysaccharides, which are deposited in the cell wall at a level of 30 ± 5%, and are closely associated with cellulose and lignin. They can be readily isolated, after delignification, by sequential alkali extraction of increasing concentration. In softwoods the predominant hemicellulose (20 ± 5%) is *O*-acetyl galactoglucomannan (Fig. 1, (**2**)). The polysaccharide is composed of a linear backbone of *β*-1,4-linked D-mannopyranosyl and *β*-1,4-linked D-glucopyranosyl residues in a ratio of approximately 4:1, respectively. Some of the mannosyl residues are substituted with (i) α-D-galactopyranosyl units at position C-6 and (ii) *O*-acetyl groups, predominantly at position C-3 and to a lesser extent at position C-2. The level of galactose sub-stitution can vary, relative to glucose, from 0.1 to 1.0. The ratio of *O*-acetyl to mannose residues is about 1:6.

An acidic arabino-4-*O*-methylgluconoxylan hemicellulose fraction is also present in softwoods at about 10%. This heteropolysaccharide is made up of a *β*-1,4-linked D-xylopyranosyl backbone substituted with α-L-arabinofuranosyl and 4-*O*-methyl-α-D-glucuronopyranosyl side groups at positions C-3 and C-2, respectively (Fig. 1, (**3**)). Generally, the xylose to arabinose to uronic acid ratio is 6:1:1. In certain softwoods, such as larch, arabinogalactan II is readily extracted by hot water to give yields of between 5% and 30%, and is a highly branched polysaccharide built up of a *β*-1,3-linked D-galactopyransoyl backbone to which at position C-6 oligosaccharide side-chains are attached.

In hardwoods, *O*-acetyl-glucuronoxylan is the predominant hemicellulose (25 ± 5%) which has the same main chain of a *β*-1,4-linked D-xylopyranosyl units, as in softwoods, and substituted with 4-*O*-methyl-α-D-glucuronopyranosyl and *O*-acetyl groups at positions C-2 and C-3/2, respectively (Fig. 1, (**4**)). Typically the xylose to acetyl to uronic acid ratio is about 10:7:1.

Furthermore, hardwoods contain a small amount of glucomannan (< 5%), which is composed of a linear chain containing *β*-1,4-linked D-mannopyranosyl and *β*-1,4-linked D-glucopyranosyl residues generally in a ratio 2:1.

3. *Lignin*

Lignin is a three-dimensional polymer built up of phenylpropane units that is laid down within the cell wall after tracheid elongation has ceased. The incorporation of lignin into the cellulose microfibril structures within the cell wall greatly enhances the mechanical strength properties of wood, over pure cellulose. Typically softwoods contain 30 ± 5% lignin while hardwoods generally have a lower lignin content (25 ± 5%). Lignin can be isolated by solvent extraction in reasonable yield from finely ground wood, after a cellulase pretreatment. Monomeric lignin precursors are *trans p*-coumaryl, coniferyl, and sinapyl alcohols (Fig. 1, (**5a**), (**5b**), and (**5c**), respectively) that undergo dehydrogenative polymerization by peroxidase and/or laccase activity to form macromolecular lignin by random coupling. The reactivity and levels of the lignin precursors govern

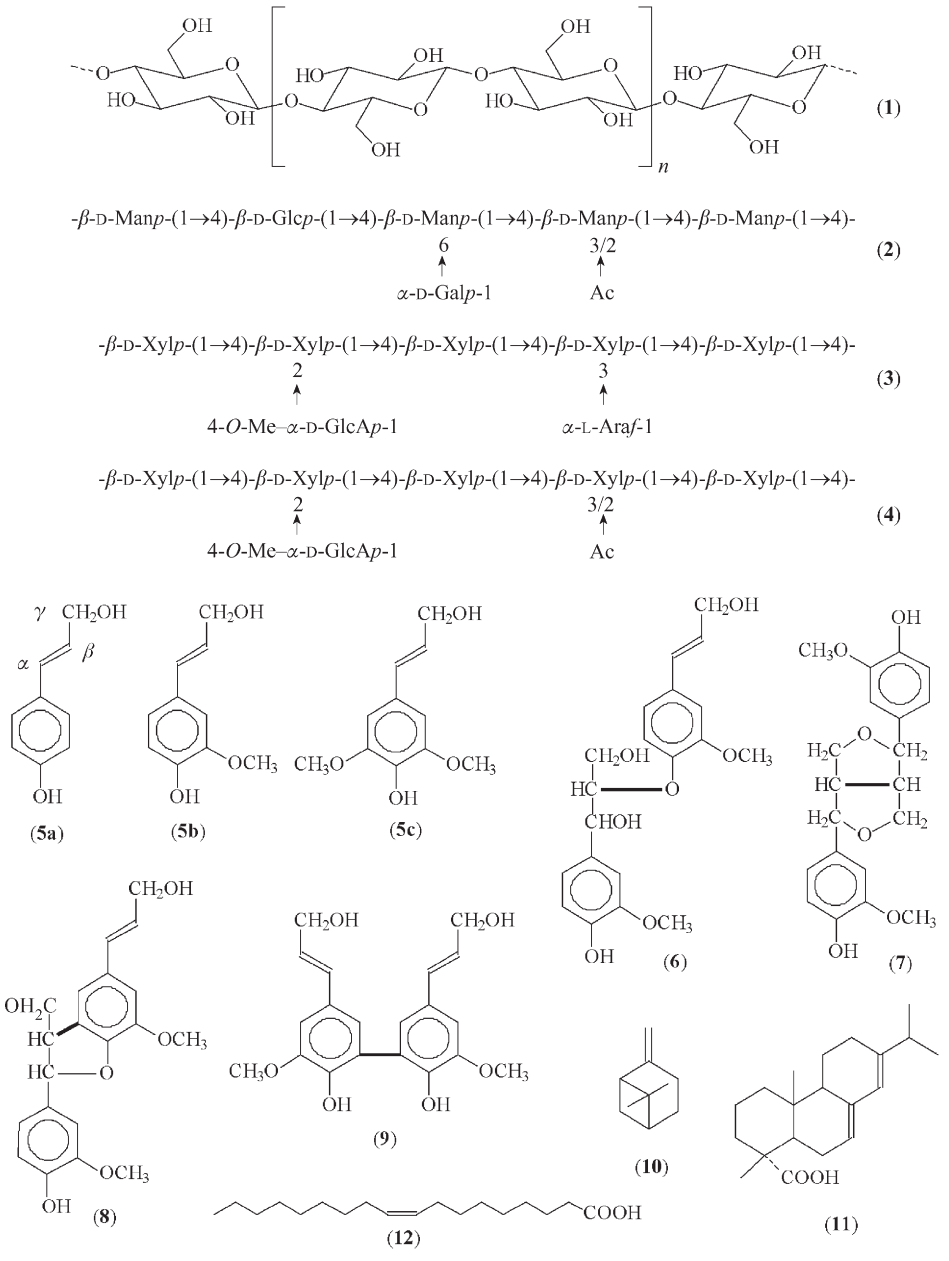

(1)
-β-D-Manp-(1→4)-β-D-Glcp-(1→4)-β-D-Manp-(1→4)-β-D-Manp-(1→4)-β-D-Manp-(1→4)-
6
α-D-Galp-1
3/2
Ac
(2)
-β-D-Xylp-(1→4)-β-D-Xylp-(1→4)-β-D-Xylp-(1→4)-β-D-Xylp-(1→4)-β-D-Xylp-(1→4)-
2
4-O-Me–α-D-GlcAp-1
3
α-L-Araf-1
(3)
-β-D-Xylp-(1→4)-β-D-Xylp-(1→4)-β-D-Xylp-(1→4)-β-D-Xylp-(1→4)-β-D-Xylp-(1→4)-
2
4-O-Me–α-D-GlcAp-1
3/2
Ac
(4)
(5a)
(5b)
(5c)
(6)
(7)
(8)
(9)
(10)
(11)
(12)

Figure 1
Common chemical structures found in wood.

the final constitution of lignin. The most predominant linkage between phenylpropane units in both softwood and hardwood lignin is the β-O-4 linkage (Fig. 1, (**6**); guaiacylglycerol-β-coniferyl ether; 50%), the β–β linkage (Fig. 1, (**7**); pinoresinol; 5%), the β-5 linkage (Fig. 1, (**8**); phenylcoumaran; 10%), and the α-O-4 linkage (10%). Softwood lignin also contains the 5–5 linkage (Fig. 1, (**9**); biphenyl; 10%). Softwoods contain lignin made up of guiacyl units, while hardwood lignin is built up from both guiacyl and syringyl units.

Production of wood pulp for the paper industry often involves the chemical removal of lignin, which is burned to drive the chemical recovery systems. A few lignin-derived chemicals are produced, such as vanillin and lignosulfonates.

4. *Extractives*

The term "extractives" means a mixture of compounds that are extracted from wood by a solvent or solvents. Lipophilic extractives are referred to as wood resin. Typically, extractives are of low molecular weight and are located outside the cell wall. Extractives, in part, contribute to the natural durability of wood by containing substances that are toxic to bacteria, fungi, and insects. In pulping, extractives can be collected as by-products such as tall oil and turpentine. However, extractives can have a negative effect on paper quality such as pitch deposits. Heartwood generally contains higher levels of extractives than sapwood. Extractives can be broken down by classes of compounds, such as monoterpenes (e.g. β-pinene; Fig. 1, (**10**)), resin acids (e.g. abietic acid; Fig. 1, (**11**)), fatty acids (e.g. oleic acid; Fig. 1, (**12**)), fats, sterol esters, phenolics, tannins, flavonoids, lignans, cyclitols, carbohydrates, and amino acids.

5. *Reaction Wood*

Reaction wood is produced as a response to stem lean, and has the function in the tree to generate growth stress to correct the stem lean, thus maintaining the most efficient position of the stem for photosynthesis. Reaction wood occurs in two different forms, compression wood in softwoods, and tension wood in hardwoods. Compression wood typically has more lignin, and less cellulose and galactoglucomannan, compared to normal wood, while tension wood has a higher cellulose content. Compression wood lignin has a higher content of *p*-coumaryl units. Both compression wood and tension wood contains a β-1,4-linked D-galactan not found in normal wood. The increased lignification of compression wood tracheids is thought to produce significant longitudinal swelling during wood formation, thus generating an expansion on the lower side of the stem, which attempts to correct the lean. In tension wood, the additional cellulose is thought to generate tension on the upper side of the stem due to shrinkage during wood formation. Unfortunately the growth stress which is beneficial to the tree has undesirable effects on timber properties resulting in abnormal shrinkage and dimensional instability.

6. *Topochemistry*

The wood cell wall has a complex structure with variations in chemical composition among different cell wall layers. Wood also contains different cell types, which in turn may have a different chemical composition. In softwoods, the tracheid accounts for 95% of the wood substance and thus has a dominant effect on the chemical composition. The tracheid cell wall consists of a highly lignified middle lamella incorporating the primary wall. The secondary wall is lignified to a much lesser extent than the middle lamella with higher lignin concentrations often present in the S3 layer lining the lumen. Both the inner S3 and outer S1 layers have a denser hemicellulose matrix than the S2 layer which is usually the thickest layer. An important effect of this complex variation in chemical composition across the cell wall is that chemical composition of a bulk wood sample can be altered significantly by changes in the proportion of the different cell wall layers. Chemical composition can also vary in proportion to wood density for the same reason. This often makes it difficult to distinguish true changes in chemical composition from changes resulting from anatomical or ultrastructural variations.

The middle lamella typically contains 20–25% of the total lignin content while the lignin concentration in this region varies from 50% to 80%. Interestingly, some of this variation has been shown to be genetically inherited among clones within a species. The secondary wall typically has a lignin concentration of 20–25%. In compression wood the distribution of lignin alters dramatically, with reduced lignification of the middle lamella and increased lignification of the secondary wall, especially in the outer S2 region at the corners of the tracheids known as the S2L region, where the lignin concentration may exceed that in the middle lamella. In tension wood the increased cellulose content is due to the presence of an additional cell wall layer consisting largely or entirely of cellulose known as the gelatinous layer because of its characteristic detached appearance under the light microscope.

In hardwoods, lignin varies in its chemical composition among cell types and cell wall regions. Guiacyl lignin predominates in cells involved with water conduction (vessels) while syringyl lignin predominates in fibers and parenchyma cells. The middle lamella region of vessels also contains guiacyl lignin while the middle lamella of fibers contains a mixed guiacyl/syringyl lignin.

7. *Conclusions*

The chemical composition of wood is highly variable and this variability has a significant effect on wood properties. Research is increasingly being aimed at the manipulation of chemical composition by genetic engineering and the use of molecular biology to study the process of wood formation. The driving force behind this research is the need to develop environmentally friendly processing options and to develop novel wood based products.

See also: Cellulose: Chemistry and Technology$^{\#}$; Wood: Macroscopic Anatomy*; Wood: Ultrastructure*

Bibliography

Fengel D, Wegner G 1989 *Wood: Chemistry Ultrastructure, Reactions.* De Gruyter, Berlin

Higuchi T 1997 *Biochemistry and Molecular Biology of Wood.* Springer, Berlin

Timell T E 1986 *Compression Wood in Gymnosperms.* Springer, Berlin, Vol. 1

A. G. McDonald and L. A. Donaldson
Forest Research, Rotorua, New Zealand

Wood: Creep and Creep Rupture

Because of the time-dependent properties of its constituents, wood experiences creep, that is, deformation resulting from constant load over time. An everyday example is the gradual sagging of a heavily loaded bookshelf. If stresses are low, the rate of creep can become negligible after 2–3 yr. At relatively high stresses, however, a deformation will be reached where the rate of creep starts to increase. Once this occurs, failure (creep-rupture) is inevitable under that load. While the underlying biological mechanisms controlling the extent of creep are complex, the engineers have to accommodate creep and creep-rupture in the design of timber structures. In some wood design codes, to avoid excessive deflections, creep effects are accounted for in calculation of deflections for the expected service life of the structure. To avoid creep-rupture of a structural member, design codes require that short-term strength values be adjusted to account for the duration of load (DOL) effect.

1. Stages of Creep

At low stress levels ($<35\%$ of the ultimate short-term strength), and under indoor environments, wood can be considered as a linear, viscoelastic material. In these cases, the creep deformation under sustained loading can be divided into two stages: a primary stage where the creep rate decreases; and a secondary stage where the creep rate is nearly constant. At higher stress levels, in addition to the primary and secondary stages, a tertiary stage, where the creep rate increases, will occur and lead to eventual creep-rupture.

Creep in wood at a given time may be described as the sum of three components: the instantaneous elastic and recoverable deformation at time zero; a delayed and recoverable elastic deformation; and a viscous deformation, which is permanent and irrecoverable. This simple description, however, may not be applicable to some wood composites. On unloading, a progressive recovery, total or partial, to the initial state occurs. Total recovery corresponds to pure viscoelasticity. Partial recovery means that plastic damage, even microscopic, has occurred.

In a stable environment, and for stresses less than about 35% of the short-term strength of the wood, the deformations under two different stresses have a constant ratio independent of time (STEP 1 1995). As a first approximation, creep compliance (strain/stress) is independent of the stress, and linear theory will be applicable. This assumption of linearity is generally appropriate at design load levels where Boltzmann's principle of superposition will be valid. However, at higher stress levels wood shows nonlinear viscoelastic behavior and Boltzmann's principle of superposition will not be applicable.

2. Factors Affecting Creep and Creep-rupture

The magnitude of creep depends on many factors such as level and type of stress, rate and duration of stress, moisture content, variations in moisture content, temperature, and grain orientation. Although there are complex interactions among these variables, and despite the fact that these interactions change from one wood material to another, they are accounted for independently in most design codes, with the exception that the combined effects of load duration and moisture content are taken into account in the European timber design code (ENV 1994).

The polymeric nature of wood constituents makes wood sensitive to moisture content and temperature. Further details about the rheological characteristics of wood and wood-based materials including a discussion on deformations at the molecular level may be found in Bodig and Jayne (1982) and in Morlier (1992). For practical purposes, the higher the moisture content and temperature, the greater the creep. However, the influence of temperature on creep is considered negligible when the temperature does not exceed 50 °C (STEP 1 1995). Higher temperatures are reported to reduce the stiffness of timber in bending, tension or compression as lignin alters its structure and hemicelluloses begin to soften (Morlier 1992). Moisture in wood acts as a plasticizer, therefore, creep increases with moisture content. The effect resulting from the interaction of moisture movement (drying and wetting) with the mechanical properties of wood is called the "mechano-sorptive effect." Fluctuations in humidity may result in increases in creep amplitude. Armstrong and Kingston (1962) report that during all desorptions, the creep increases faster than it would at constant moisture content, regardless of the rate of desorption. This phenomenon also occurs in the first sorption, but during subsequent sorptions at the same moisture level, the creep deflection usually decreases.

Creep increases with load duration. It strongly depends on load level, which must be lower when the duration is extended, in order to ensure an acceptable deformation limit with time. If the duration of applied stress is sufficient, wood may commence tertiary creep preceding rupture, which can be described as progressive damage in the wood. The higher the stress, the shorter the time to rupture. In other words, the time to creep-rupture depends on the sustained load level. In design codes, the specified strength is selected such that creep remains within the secondary stage where the rate of deformation is low and stable during the lifetime of the structure. Experimental results show that this stability exists at stress levels below 35% of the short-term strength.

Creep deflections also occur in other types of loading, e.g., a ramp load may be considered as a series of step loads for which Boltzmann's principle of superposition may be used to predict creep. When the loads are increasing with time, wood exhibits the influence of rate of loading by showing relatively larger strain under slowly increasing stress. Thus, specimens tested at a slower rate of loading will fail at a lower stress level (Karacabeyli and Barrett 1993).

3. *Modeling Techniques*

Modeling of the creep deformation behavior of wood has followed the outlines suggested by Alfrey (1948). The mathematical models here used to describe linear viscoelastic behavior resemble models built from discrete elastic and viscous elements. To adequately describe wood, multi-parameter Voight or Maxwell models have been used. These models allow calculation of the time-dependent uniaxial stress-strain behavior for many types of load excitations such as: creep (sustained loading), stress-relaxation (constant kept deformation), constant rate of ramp loading, and vibration-damping behavior at various frequencies.

To account for creep in design, engineers rely on good judgment supported by available models. Wood researchers applied a variety of techniques in modeling creep behavior (Schniewind 1968, Bach and Pentoney 1968, Morlier 1992, Bodig and Jayne 1982, Fridley *et al.* 1992). Models based on damage accumulation (Foschi *et al.* 1989) and fracture mechanics (Valentin *et al.* 1991) also have been used in predicting creep-rupture of wood.

4. *Evaluation of Creep and Load Duration Performance*

Creep deformation and DOL strength effects on lumber and some panel products have been studied widely, and research findings have been implemented in various design codes. For example, in the European timber design code (ENV 1994), material-specific creep and load duration factors are specified for several load duration and service classes. This format allows the factors to reflect the interaction between load duration and moisture effects.

Given that there may be substantial differences between the long-term structural performance of different composite wood products, it is important to determine the creep and DOL performance for each structural composite wood product. This may be achieved by showing an equivalency to lumber through a verification program, or by quantifying the factors through a comprehensive testing and reliability study.

4.1 *Verification Method*

A verification approach is a pass/fail method, which requires the producers to demonstrate the equivalency of the long-term structural performance of their products to that of lumber. The following method has been successfully used for a number of engineered wood products in North America for indoor environments. In this method, a representative sample is subjected to a constant load for a time longer than the expected duration of extreme occupational and snow loads. In other words, the level of the constant load tests is selected to represent overload levels. Usually a minimum sample size of 28 specimens is sampled and tested for 3–4 months. The constant load level is selected as a percentage of the short-term strength of a matched control group, determined in ramp tests of duration from 1–10 min. A constant load level of 55% of the lower fifth-percentile point estimate came into use in the 1990s in North America. Finally, an equivalency of the creep and creep-rupture behavior of a composite wood product to lumber can be shown by comparing some of its performance parameters to those established for lumber.

The following sections describe the creep and DOL performance parameters used as criteria in the verification method for the evaluation of the long-term structural performance of a wood product. If the product satisfies all the criteria, then it passes the verification test and designers may use the creep and load duration factors assigned for lumber for that product.

(*a*) *Stress levels for avoiding creep-rupture*. Under certain constant stress levels, lumber does not fail within a given time, e.g., the large database developed under the US/Canada Duration of Load Test Program (Karacabeyli and Soltis 1991) shows that lumber does not fail under a stress level of 55% of its short-term strength within 3 months under bending, tension or compression parallel-to-grain loading. The equivalent load level for 2 years is about 45%. Laufenberg *et al.* (1999) show that plywood and oriented strandboard (OSB) also will not fail over a 3-month period when subjected to a constant bending stress level of 55%. For loads perpendicular to grain, this ratio may be different. Madsen (1972) reports that for various rates of tension loading perpendicular to the grain, the reduction in strength is considerable with decreasing rate of loading. Bach (1975) shows that the effect of sustained loading on relative strength is more severe in bending perpendicular to grain than in bending parallel to the grain. Three-month constant bending load tests at 55% stress level have come into use in North America for the creep and creep-rupture evaluation of engineered wood products.

(*b*) *Decreasing creep rate.* A criterion based on the decreasing creep rate is used to ensure that no individual specimen shows evidence of tertiary creep that signals an eventual failure. For example, creep in the first month is usually significantly greater than that in the second month, and second month creep is slightly greater than or equal to third month creep. This suggests that the specimen is exhibiting primary creep in the first month, and secondary creep in the second and third months. If the creep rate increases due to fluctuations in the environmental conditions, the test period is normally extended for a month to show that the increase is not due to the onset of tertiary creep.

(*c*) *Fractional deflection limit.* If a wood product satisfies the above two criteria but exhibits relatively high absolute creep deflections, a criterion based on fractional deflection is being used, that is, the ratio of total deflection and initial deflection, where total deflection is the sum of initial deflection and creep deflection. A fractional deflection requirement of 2.0 (i.e., a relative creep of 1.0) established from lumber data is used as the fractional deflection criterion.

(*d*) *Permanent deformation limit.* Although generally a one-month creep-recovery period is used to determine the permanent deformations, the recovery in fact continues for almost three times the constant load period. Naturally, most of the recovery happens in the first month. After the creep recovery period, the ratio of the permanent deflection and creep deflection can be determined. This ratio is a good indicator of damage, e.g., a ratio of zero indicates that the entire creep was due to delayed elastic deformations, and is fully recovered (purely viscoelastic) when the specimen is unloaded. Alternatively, a ratio of unity indicates that the entire creep deflection is nonrecoverable, viscous deflection. The creep-recovery test allows us to check if all the individual specimens show delayed elastic deflections, and if nonrecoverable viscous deformations are within acceptable levels after the removal of the load. The author has evaluated several laminated veneer lumber (LVL) products under a 3-month constant bending load at 55% of the fifth percentile point estimate of the control group. Within a month of creep recovery, all individual specimens showed delayed elastic recovery. On average, these products recovered about 30% of the creep deflection during 1 month of creep recovery.

(*e*) *Residual stiffness and strength.* Predicting long-term properties of a new product from a three-month constant load test has limitations. An assessment of residual stiffness and strength may provide the engineers with increased confidence. The author, therefore, recommends that after the completion of creep and creep-recovery tests, the survivors be tested under short-term ramp loading with the same loading rate used for the control group. By collecting strength and stiffness data before and after the constant load tests, the strength and stiffness degradation due to long-term loading can be assessed. Several LVL products that the author evaluated did not exhibit any appreciable degradation when subjected to a constant bending load of 55% of the fifth-percentile point estimate of the short-term strength of the control group. Based on the results obtained to date using that stress level, a criterion based on an average 5% maximum degradation for the group may be used both for strength and stiffness.

The above five criteria may be used to evaluate product's equivalency in indoor environments. However, if the product was going to be used in other applications, appropriate additional testing would be necessary. For example, constant-load and slow ramp loading tests under high and cyclic humidity conditions would be very useful for assessing the performance for outdoor environments. For end-use applications requiring good performance in other conditions, e.g., elevated temperatures, it is advisable to design a program of testing, which includes those conditions.

4.2 Probabilistic Approach for Quantification of Load Duration and Creep Factors

Load duration and creep factors for a particular wood product may be determined by following procedures similar to those employed for lumber (Foschi *et al.* 1989, Ellingwood and Rosowsky 1991). These procedures include substantial testing and analysis. In summary, the following are recommended for this approach:

(i) Establish a minimum of five matched groups, one for short-term (control) testing, two for creep testing, and two for creep-rupture testing. Ramp test the control group to failure within 10min. (under constant rate of loading or displacement), and obtain short-term strength distribution. Test the creep groups under two relatively low constant stress levels (e.g., 15% and 30% of the mean strength of the control group) for 6 months. Unload the creep groups for 1 month for creep-recovery data. Test the two creep-rupture test groups under relatively high constant load levels (e.g., 60 and 75% of the mean strength of the control group) for 1 year. Fit a damage accumulation model to the time-to-failure data.

(ii) By using selected load distributions and the short-term strength distribution, develop reliability indices without DOL effects for different performance factors. Repeat the same analysis with the inclusion of a damage accumulation model to account for the load duration effects. The DOL can then be determined as the ratio of two performance factors obtained from those two cases, at a given reliability index. Creep factors may be developed by using the relative creep values, or by carrying out a reliability analysis similar to that described above, but using a creep model.

5. Test Considerations

The following parameters should be considered for creep and creep-rupture behavior: (i) intended application of the product which should be reflected in test conditions, service history, and experience; (ii) material size, span, grade, species, quality, and class; (iii) load type (ramp, constant, cyclic), mode (bending, shear, tension, compression, parallel or perpendicular-to-the-grain) and level; and (iv) process variables. Matching between constant-load test groups and control groups requires special care, e.g., with composite products, pair matching (cutting lengthwise into two pieces—test and control) has proven to be a reasonable method. Conditioning must be done in the same environment.

5.1 Short-term (Control) Testing

Usually, short-term tests and long-term tests are performed in different test machines, which may cause different results, e.g., if bending tests are used, the lateral supports in both bending test machines should be similar.

5.2 Long-term Testing

There are many designs of testing frames for constant load tests. Detailed descriptions of some of these may be found in SP-27 (Forintek Canada 1986).

6. Concluding Remarks

Creep and creep-rupture in wood are important parameters that should be considered in the design of timber structures. As described here, creep is a combination of elastic and viscous deformations that occurs in three stages. It is most strongly affected, in normal environments, by load level, load type, load duration, and moisture movement (drying and wetting); whereas temperature has a significant effect above 50 °C. Two methods are presented that can be employed to evaluate the creep and creep-rupture performance of wood-based products: a pass/fail verification method which requires producers to demonstrate equivalency of the creep and creep-rupture performance of their products to that of lumber, and a probabilistic method, useful for quantification of design factors for creep and creep-rupture. Finally, short- and long-term test considerations are discussed.

See also: Wooden Structures$^{\#}$; Wood: Strength and Stiffness*; Paper: Creep*

Bibliography

Alfrey T 1948 *Mechanical Behavior of High Polymers*. Interscience, New York

Armstong L, Kingston R 1962 The effect of moisture content changes on creep in wood. *Nature* **185**, 862–3

Bach L, Pentoney R E 1968 Non-linear mechanical behavior of wood. *For. Prod. J.* **18** (3), 60–6

Bach L 1975 Failure perpendicular to the grain in wood subjected to sustained loads. *Wood Sci.* **7** (4), 323–7

Bodig J, Jayne B A 1982 *Mechanics of Wood and Wood Composites*. Van Nostrand Reinhold, New York

Ellingwood B, Rosowsky D 1991 Duration of load effects in LRFD for wood construction. ASCE. *J. Struct. Eng.* **117** (2), 584–99

ENV 1994 Eurocode 5, Design of Timber Structures, Part 1–1. British Standards Institution, London

Forintek Canada 1986 *Proc. Int. Workshop Duration of Load in Lumber and Wood Products*. Special Publication No. SP-27. Forintek Canada Corp., Vancouver, Canada

Foschi R O, Folz B R, Yao F Z 1989 *Reliability-based Design of Wood Structures*. Structural Research Series, Report No. 34. University of British Columbia, Vancouver, Canada

Fridley K J, Tang R C, Soltis L A 1992 Load-duration effects in structural lumber: strain energy approach. ASCE. *J. Struct. Eng.* **118** (9), 2351–69

Karacabeyli E, Barrett J D 1993 Rate of loading effects on strength of lumber. *For. Prod. J.* **43** (5), 28–36

Karacabeyli E, Soltis L A 1991 State-of-the-art report on duration of load research for lumber in North America. In: Marcroft J (ed.) *Proc. Int. Timber Engineering Conf.* London, Timber Research and Development Association, pp. 4141–55

Laufenberg T L, Palka L C, McNatt J D 1999 Creep and creep-rupture behavior of wood-based structural panels. Research Paper FPL-RP-574 USDA Forest Products Laboratory, Madison, WI

Madsen B 1972 Duration of load test for wood in tension perpendicular-to-the-grain. SRS Report. 7. University of British Columbia, Vancouver, Canada

Morlier P (ed.) 1992 Creep in timber structures. Report of RILEM Technical Committee 112-TSC. RILEM (The International Union of Testing and Research Laboratories for Materials and Structures). E & F N Spon, London

Schniewind A P 1968 Recent progress in the study of rheology of wood. *Wood Sci. Technol.* **2**, 188–206

STEP 1 1995 *Timber Engineering STEP* 1. Centrum Hout, The Netherlands

Valentin G H, Bostom L, Gustafson P J, Ranta-Maunus A, Gowta S. 1991 Application of fracture mechanics to timber structures, RILEM state-of-the-art report. Research Notes 1262. Technical Research Centre of Finland

E. Karacabeyli

Forintek Canada Corp., Vancouver, British Columbia, Canada

Wood: Macroscopic Anatomy

Wood is a cellular material produced by a living tree to support the leafy crown, conduct water and dissolved nutrients from the roots upward to the crown, store reserve food (primarily carbohydrates) that is later used for tree growth, and to synthesize extractives during heartwood formation or in response to injury. The anatomy of wood reflects these biological functions. Support and conducting cells are usually "pipe-like"; they have rigid walls, generally are longer than they are wide, and normally are oriented with their long axes parallel to the long axis of the tree. Water-conducting cells are dead hollow cells, as are the support cells in most commercially important US woods. Storage cells (parenchyma) generally are very small, and are aggregated into horizontally oriented structures called rays. In addition, in many hardwoods and a few softwoods there may be longitudinally oriented storage cells. Reserve food materials are moved "in and out" of the wood via the rays. The anatomy of wood affects its material characteristics. For example, the interfaces between rays and the longitudinal elements seem to be "weak points" in wood. It is at these interfaces and within the rays that separations between cells (checks) originate during wood drying, and wood is easiest to split.

As one would expect for a substance of biological origin, wood is a variable material. The types of woody cells, their relative abundance, and arrangements differ among species. There also is within species a variation in cell structure related to growing conditions, and within trees variation related to tree age. Major challenges for the forest products industry have been, and continue to be, the ability to grow wood that is predictable in its behavior and to match wood characteristics to product requirements.

1. Wood Types: Hardwoods and Softwoods

In the forest products industry, the terms softwood and hardwood are applied to two major groups of trees that correspond to two major botanical groupings. Softwood trees have seeds that are not covered, but are produced in cones (conifers such as pine, redwood, cedar, cypress). Hardwood trees produce covered seeds within flowers (dicotyledonous angiosperms such as oak, ash, maple, teak, mahogany). The use of the terms softwood and hardwood is somewhat unfortunate as it implies absolute differences between the two groups in the hardness of their woods. However, the wood of some hardwoods is softer than that of some softwoods (and vice versa). One of the world's lightest woods, *Aeschynomene hispida* (30 kg m^{-3}), comes from a dicotyledonous tree, and, in that sense, is a hardwood. For commercially important woods native to the USA the density of softwoods is 290–600 kg m^{-3}, average 410 kg m^{-3}; of hardwoods is 320–810 kg m^{-3}, average 500 kg m^{-3}. Density is more variable within tropical woods than it is within temperate-zone woods.

The wood structure of hardwoods and softwoods is distinct, with softwoods having the simpler wood anatomy. In softwoods, a single cell type, the longitudinal tracheid, functions for both support and conduction, and constitutes 90% of the wood volume (Fig. 1). Longitudinal tracheids are usually 100 times as long as they are wide, with lengths usually of 2–4mm, and up to 8mm in large softwood trees, such as redwood (*Sequoia*). Rays in softwoods account for no more than 10% of the wood volume, and are very narrow.

Hardwoods have vessels for water conduction, and fibers for support (Fig. 2). Vessels are composite structures composed of open-ended vessel elements that have relatively large diameters, thin walls, and large central openings (lumens). Individual vessel

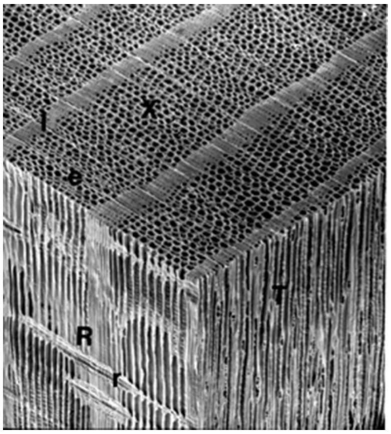


Figure 1
Scanning electron micrograph of softwood block: X, cross-section; e, earlywood with longitudinal tracheids with relatively wide openings; l, latewood with longitudinal tracheids with narrow openings; R, radial section; r, ray extending across growth ring boundaries; T, tangential section (courtesy of N. C. Brown Center for Ultrastructure Studies, SUNY College of Environmental Science and Forestry).

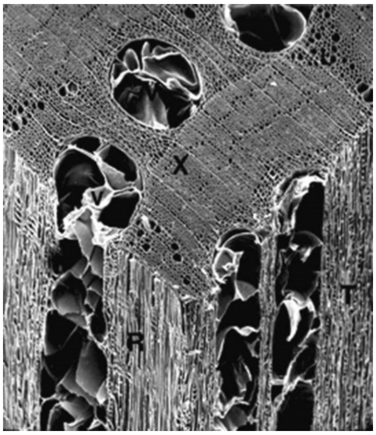


Figure 2
Scanning electron micrograph of ring porous hardwood block: X, cross-section; T, tangential section; R, radial section; V, vessel. Note the wide vessels (V) that are filled with tyloses. In the conducting sapwood the vessels would be hollow. In this wood, the tyloses fill the vessels and reduce permeability (courtesy of N. C. Brown Center for Ultrastructure Studies, SUNY College of Environmental Science and Forestry).

elements are usually 25–300 μm wide and 0.2–1mm long. The composite vessels range from a few centimeters to more than 10m long. Fibers are narrow tubes (usually 10–30 μm wide and 1–2 mm long) with varying cell wall thickness. Support and conducting cells usually are dead at functional maturity, although in some species fibers are living in the sapwood and are involved in food storage. Ray size varies between hardwood species. In commercially important hardwoods, rays usually account for 10–30% of the wood volume, and vary from 1 cell wide and not visible to the eye (e.g., willow (*Salix* spp.) and cottonwood (*Populus* spp.)), to more than 20 cells wide and visible on all surfaces of the wood (as in oak (*Quercus* spp.)).

Proportions and arrangements of the different cell types, as well as cell wall thickness, differ from species to species. Woods that have a high proportion of thick-walled fibers are hard and heavy. Hickory (*Carya* spp.) is such a wood, and is the preferred material for tool handles in the USA. Woods with a high proportion of vessels and/or thin-walled fibers are soft and light. Balsawood (*Ochroma* spp.) is such a wood, and is used for model airplanes, and as an insulator because of the high proportion of open lumens.

2. *Cell Orientation*

The orientation of the woody cells within the trunk is of major importance in explaining many of the properties of wood. Cells that function in support and conduction, and, at times, some parenchyma cells, have a longitudinal orientation, i.e., the long axes of the cells are parallel to the long axis of the tree. Grain refers to the orientation of the support and conduction cells. In straight-grained wood, the support and conducting cells are oriented parallel to the outside of the tree; in spiral-grained wood, the support and conducting cells literally spiral (right-handed, left-handed, or alternating between the two) around the tree trunk. Some species (e.g., sweetgum (*Liquidambar styraciflua*)) are more apt to produce spiral grain than others. Straight-grained wood is preferred for most applications, as it is easier to dry without warping and twisting. However, in some mahogany trees, alternating spiral grain imparts a pleasing appearance to the wood, and such ribbon-striped mahogany is valued as a decorative veneer and for inlays.

Because wood is composed of cells oriented longitudinally and horizontally, it will have a different appearance depending on how it is cut, as different aspects of the cells are exposed (Fig. 3). A cross-section is produced when the cut is at right angles to the long axis of the tree. When a tree is cut down, the end of the stump reveals a cross-section. In this section, growth rings are evident as a series of concentric circles, and rays appear as fine lines crossing the growth rings at right angles. Longitudinal sections are produced by cutting parallel to the long axis of the tree. There are two types of longitudinal sections: radial and tangential. When the longitudinal cut is made by cutting parallel to the rays and perpendicular to the growth rings, a radial section is produced. In radial sections rays are viewed from the side, look like stripes, and have a "brick-wall" appearance with the individual cells in a ray being equivalent to individual bricks in the wall. Hard maples and oaks have a distinctive ray fleck on their radial surfaces. A tangential section is produced by cutting parallel to growth ring boundaries, or the surface of the tree, and cutting across the rays at right angles. The ends of rays are visible in tangential sections; ray height and width can be determined from tangential sections. Lumber cut along the radial section is termed quarter-sawn and that cut in the tangential plane is flat-sawn.

Wood behaves differently along the different axes. Water is lost most readily from the cross-section, and penetration of fluids (e.g., preservatives, pulping liquor) is easiest into the cross section. It is easiest to split wood longitudinally, particularly along rays, as it is easier to split cells apart, rather than cut across their cell walls. Most of the mechanical properties of wood are greatest along the grain, while shrinkage and swelling with changes in moisture content are less along the grain.

3. *Growth Rings*

New wood and new inner bark are inserted between existing wood and inner bark by cell divisions within

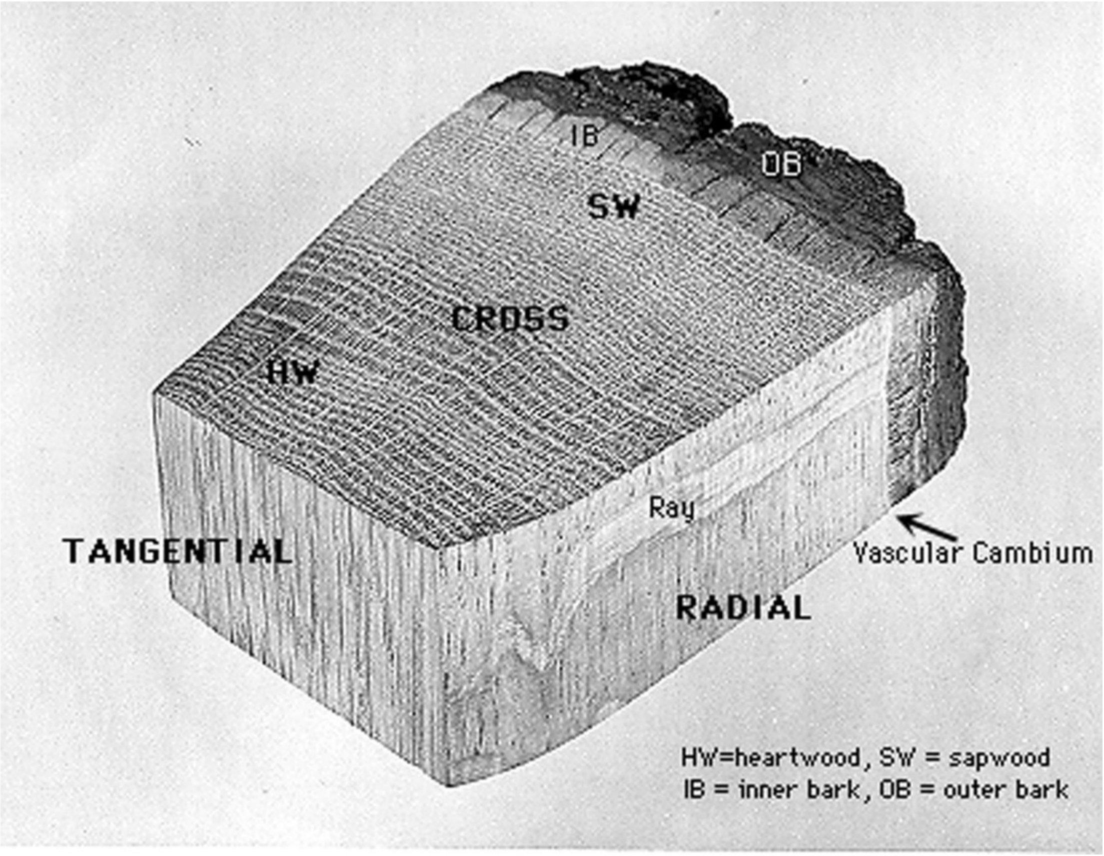


Figure 3
Block of red oak (*Quercus* sp.) showing bark (OB, outer bark composed of alternating layers of cork and dead food-conducting cells; IB, inner bark with functional food-conducting cells) and wood (SW, sapwood, in oak usually 10–12 outer rings; HW, heartwood, the region where the storage cells are dead). Cross, radial, and tangential surfaces are labeled as such, and the broad rays of oak are visible in all three surfaces.

the vascular cambium, a layer that is considered to be but one cell wide. Stripping the bark from around a tree can seriously injure (or kill) it, as usually the vascular cambium will also be stripped off, and the ability of the tree to form new conducting tissues is seriously compromised.

When trees grow in seasonal climates, they almost always produce trunk wood with growth rings. Cell production varies with the season, and there often are differences in the texture of the wood formed during different seasons. The wood formed at the beginning of the growing season is the earlywood (in north temperate regions often referred to as springwood); the wood formed later in the season is the latewood (in north temperate regions often referred to as summerwood). The latewood generally is denser and darker than the earlywood. Growth rings are visible because of the difference in texture between the latewood (usually comprised of relatively small and thicker-walled cells) and the earlywood of the subsequent year (with relatively large and thin-walled cells). In the north temperate region, a growth ring usually represents the amount of wood accumulated during one year and so is often called an annual ring. In the tropics, growth rings can be distinct, indistinct, or absent. If growth rings are present, they do not necessarily represent the growth of a single year, since growth often reflects wet and dry seasons. Determining the age and growth rate of tropical trees is more difficult than for temperate woods because growth rings do not always correspond to annual rings.

The appearance and properties of wood are affected by the ratio of earlywood to latewood, and how well defined the differences are in the transition between these zones. Diffuse porous hardwoods do not have a pronounced difference between earlywood and latewood, and the size and spacing of the vessels is nearly the same throughout a growth ring (Fig. 4).

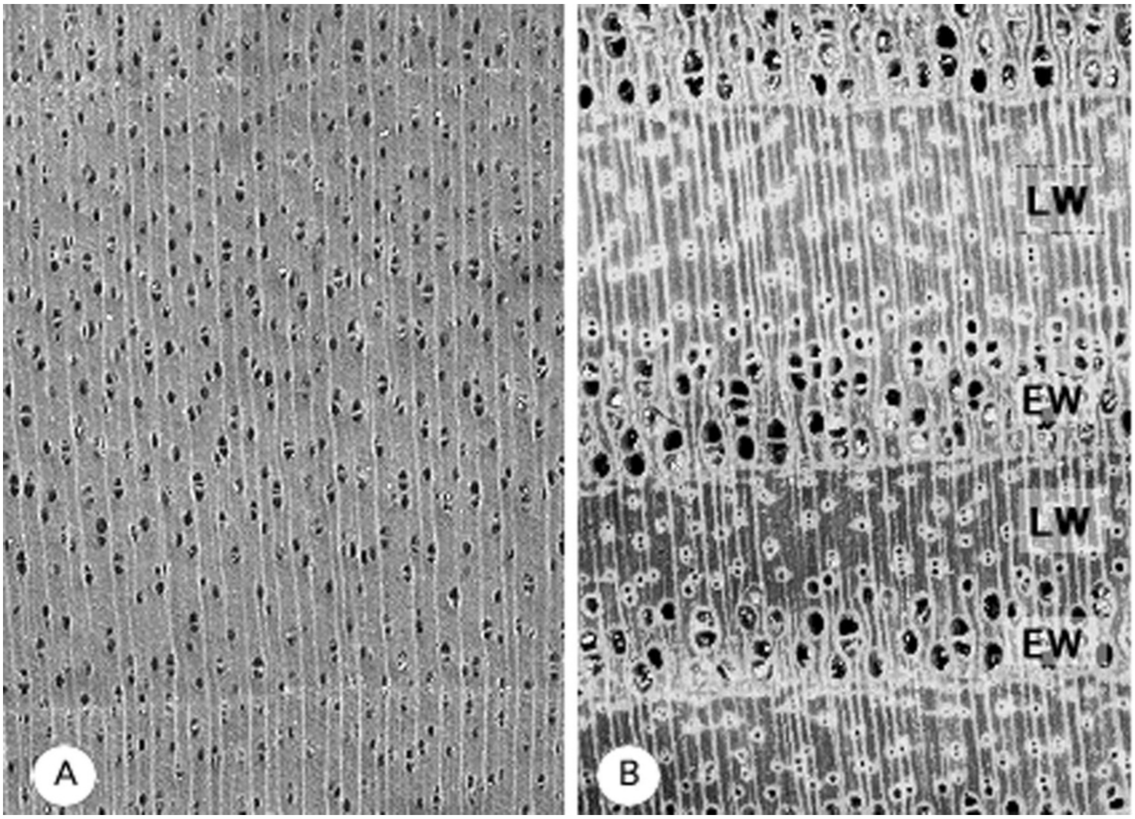


Figure 4
(A) Birch (*Betula* sp.), an example of a diffuse porous hardwood; vessel size and distribution is similar throughout the growth ring. (B) Ash (*Fraxinus* sp.), an example of a ring porous hardwood. The earlywood zone (EW) has large diameter vessels and the latewood (LW) has narrow vessels. Earlywood width does not vary much from year to year, but latewood width varies depending on growing conditions.

Almost all tropical hardwoods are diffuse porous. Ring porous hardwoods have earlywood with large diameter vessels, and a distinct latewood zone with small diameter vessels (Fig. 4). Approximately 25% of the hardwoods from the north temperate zone are ring porous. Some woods are intermediate between diffuse porous and ring porous and have a gradual change in vessel diameter throughout the growth ring (e.g., walnut (*Juglans* sp.)), and are called semi-ring porous.

Growth rate affects the density and openness of ring porous woods (e.g., oak (*Quercus* spp.), ash (*Fraxinus* spp.), hickory) more than it affects diffuse porous woods (e.g., birch (*Betula* spp.), maple (*Acer* spp.)). In ring porous woods, within a growth ring there is an earlywood zone that is of near-constant width each year, while the width of the latewood zone varies. Slow-grown ring porous woods with narrow growth rings will have proportionally more earlywood with a high vessel volume than latewood per unit area, and so will have a lower density than rapid-grown ring porous woods that will have proportionally more latewood with a high fiber volume and, consequently, a higher density.

Softwoods also vary in the proportions of earlywood and latewood, and the degree of difference between the two regions. Some species have an abrupt transition from the earlywood to the latewood, and a distinct band of latewood is formed (e.g., hard pines such as loblolly pine (*Pinus taeda*)). Others (e.g., soft pines such as eastern white pine (*P. strobus*)) have a gradual transition from earlywood to latewood, and not much latewood at all. The relationship between ring width and density in softwoods varies; in some species growth rate does not affect density, while in others an increase in growth rate can result in either a decrease or an increase in density.

Cell size and variation therein affect wood texture. Hardwoods with narrow vessels and fibers, and softwoods with narrow tracheids are fine-textured; hardwoods with large diameter vessels, and softwoods with wide tracheids are coarse-textured. The hardwoods boxwood (*Buxus* sp.) and dogwood (*Cornus florida* L.) and the softwood yew (*Taxus* spp.) are noted for their even, fine texture and even wear during use. Ring porous hardwoods and softwoods with a distinct latewood band have an uneven texture, because of the pronounced difference in earlywood and latewood cell sizes. Some textures are correlated with particular regions of the world; hardwood trees with diffuse porous wood and large diameter vessels (more than 200 µm across) only occur in the tropics.

4. Sapwood and Heartwood

The outer wood with living parenchyma cells is sapwood. The number of years that the parenchyma cells live and therefore the width of the sapwood zone varies from species to species, but parenchyma cells eventually die, and the wood with only dead cells is termed heartwood. Often, but by no means always, the heartwood is darker colored than the sapwood. Before dying, parenchyma cells often manufacture extractives. These extra chemicals may color the wood, coat cell walls and passages between cells and thus lower permeability, change the affinity of the wood for water, alter its dimensional stability, and affect its durability (resistance to decay). For instance, heartwoods of redwood, cedar, and teak are noted for their durability and are used for outdoor furniture and shingles, but the sapwoods of redwood and cedars with very low extractive content are not resistant to decay.

4.1 Tyloses

During normal heartwood formation or after an injury, parenchyma cells adjacent to a vessel may "grow into" that vessel. This occurs usually in hardwood species where the interconnections between the vessels and the storage cells (ray parenchyma) are large. These infillings of the vessels by parenchyma cells are called tyloses, and they lower the permeability of the wood. Woods with tyloses are slower drying and more difficult to treat with preservatives than woods without tyloses. White oak is noted for having many tyloses and a very low permeability.

5. Reaction Wood

It is termed abnormal wood, but the ability to form reaction wood to maintain the vertical alignment of a tree is important for the long-term survival of tree species. Reaction wood is of widespread occurrence. In hardwoods, reaction wood is formed on the upper side of a bend or lean and so is termed tension wood. If the growth rings are eccentric it is likely that tension wood is present. Tension wood often is difficult to recognize without microscopic examination, and can be found in straight stems. Tension wood fibers have a different wall structure than normal wood fibers. This variation is correlated with tension wood twisting and warping during drying, and being difficult to plane to a smooth surface. In extreme cases, logs with large amounts of tension wood will split immediately upon being sawn.

In softwoods, reaction wood is compression wood, and formed on the underside of a bend or lean. Compression wood also warps during drying, and is brash, breaking suddenly under load. It is also less permeable than normal wood, and has shorter cells. Unfortunately, pines are prone to form compression wood. Minimizing and eliminating compression wood improves the quality of softwoods for solid wood products as well as for pulpwood.

See also: Wood, Constituents of*; Wood: Ultrastructure*

Bibliography

Haygreen J G, Bowyer J L 1997 *Forest Products and Wood Science. An Introduction*, 3rd edn. Iowa State University Press, Ames, IA

Hoadley R B 1980 *Understanding Wood. A Craftsman's Guide to Wood Technology*. Taunton Press, Newtown, CT

Jane F W (revised by Wilson K, White D J B) 1970 *The Structure of Wood*. Adam & Charles Black, London

Panshin A J, de Zeeuw C 1980 *Textbook of Wood Technology: Structure, Identification, Properties, and Uses of the Commercial Woods of the United States and Canada*, 4th edn. McGraw-Hill, New York

Wilson K, White D J B 1986 *The Anatomy of Wood: Its Diversity and Variability*. Stobart & Son, London

E. Wheeler
North Carolina State University, Raleigh, North Carolina, USA

Wood: Moisture Content, Hygroscopicity, and Sorption

Wood is a hygroscopic material that has the ability to adsorb or desorb water in response to temperature and relative humidity of the atmosphere surrounding it. This affinity of wood for water is caused by hydroxyl groups accessible in the cell walls of wood. As a consequence, the moisture content of wood is one of the most important variables affecting its physical and mechanical properties.

1. *Definitions and Types of Water in Wood*

The moisture content (M) of wood is expressed as a ratio of the mass of water (kg) per unit mass of oven-dry wood (kg):

$$M = \frac{w_i - w_o}{w_o} \times 100\% \tag{1}$$

where w_i is the initial mass and w_o is the oven-dry mass. In terms of wet basis, M is given by:

$$M = \frac{w_i - w_o}{w_i} \times 100\% \tag{2}$$

The fractional moisture content (m) is given by $m = M/100$.

The elemental composition of wood is carbon, hydrogen, and oxygen. These are combined into organic polymers, cellulose, hemicelluloses, and lignin, which are the primary components of wood (Haygreen and Bowyer 1996). The approximate percentages of these polymers making up the cell walls are 40–50%, 20–35%, and 15–35% respectively. Secondary components are tannins, volatile oils, resins, and ash (Panshin and de Zeeuw 1980). The primary components each adsorb moisture; hemicellulose absorb more compared with cellulose and lignin (Stamm 1964). They are mixed in the cell walls in a complex and highly ordered manner, which is reflected directly in the physical and mechanical properties of wood (Panshin and de Zeeuw 1980). Some species of wood contain nonwater- or water-soluble extractives and have a significantly lower hygroscopicity than other woods.

There are two types of water in wood. The "water of constitution" is the water included in the chemical structure of wood and it is inherent to the organic nature of the cell walls. It cannot be removed without modifying the chemical composition of wood. The second type of water comes in three forms: "bound" or "hygroscopic" water which is adsorbed by sorption sites in amorphous areas of cellulose and hemicellulose present in the cell walls; "free" water which is liquid-like water in the cell lumen and voids; and "vapor" water which is present in voids and cell lumen unless the wood is completely saturated.

As wood begins to lose moisture when exposed to ambient conditions in the atmosphere, water leaves from voids and cell lumen first while the bound water content remains constant. The moisture content level which corresponds to lumen containing no free water (but will contain water vapor), while no bound water has been desorbed from the cell wall material, is known as the fiber saturation point (M_F). For most temperate woods, M_F is in the range 25–30%, and is lower for tropical woods that have higher extractive contents. As the moisture content of wood decreases below M_F, bound water will begin to leave the cell wall material; most of the physical and mechanical properties vary with moisture content below M_F. At a certain point, equilibrium is attained between moisture in the wood and that in the surrounding atmospheric environment. This is where the wood will not gain or lose any moisture with time and it is known as the equilibrium moisture content (M_E) of the wood.

2. *Measurement of Moisture Content*

Several methods are used to determine the moisture content of wood. ASTM Standard D4442 outlines the accepted methods to determine the oven-dry mass of wood. The oven-dry method is the sole primary method where it has the highest accuracy or degree of precision for research purposes. A forced-convection oven is required to be maintained at a temperature of 103 ± 2 °C throughout the drying time. Specimens are weighed using a balance of the desired precision and then are placed in the oven. The endpoint is reached when the mass loss in a 3 h interval is equal to or less than twice the selected balance sensitivity. Other methods outlined in ASTM D4442 are less precise: distillation, Karl Fischer titration, infrared (heating and absorption), microwave (heating and absorption), NMR, and vacuum-oven drying.

Two types of electric moisture meters are used to measure the moisture content of wood: resistance and dielectric or power-loss meters. These meters are suitable for industrial uses as they are nondestructive to the wood and give immediate results. The resistance meter uses the d.c. resistance of wood to indicate the moisture content and is limited to the moisture content range $M = 6\%$ to M_F. The dielectric meter is based on the capacitance property of the wood and is limited to $M = 0\%$ to M_F range or higher. Other industrial moisture meters that exist, but have their limitations, are: laser, mass (density), microwave, and gamma ray.

3. *Sorption Isotherms*

A graphical representation of the equilibrium moisture content as a function of relative vapor pressure, h, or percentage relative humidity, $H = 100h$, at a constant temperature is referred to as a sorption isotherm. There are six different sorption isotherms and the wood moisture isotherm is described as Type 2. Figure 1 shows the wood moisture sorption isotherm for desorption and adsorption.

Never-dried wood, referred to as green, has a desorption isotherm higher than oven-dried wood adsorption and desorption isotherms. Also, the oven-dried desorption isotherm is always higher than the adsorption isotherm. The two desorption isotherms are the same for below 50% relative humidity. The difference between adsorption and desorption curves is known as sorption hysteresis and is present for many hygroscopic materials (Skaar 1988). It is

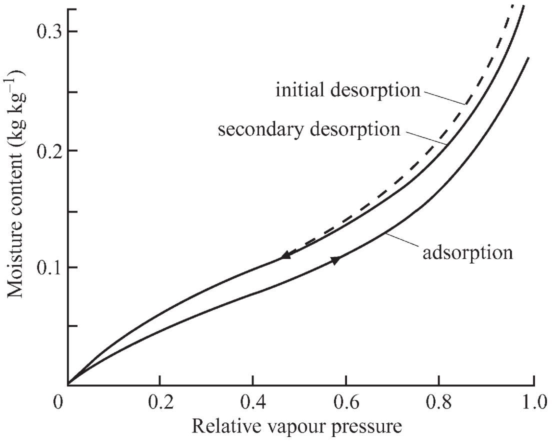


Figure 1
Typical initial desorption, adsorption, and secondary desorption moisture isotherm for wood at 25 °C.

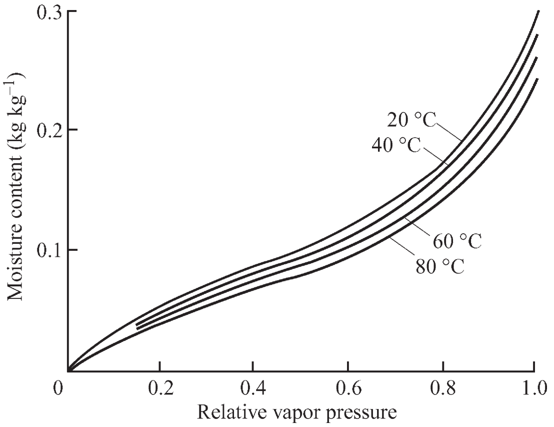


Figure 2
Effect of temperature on typical moisture isotherms for wood.

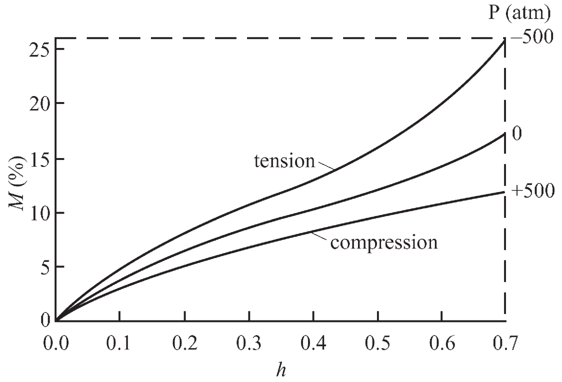


Figure 3
Sorption isotherms for the cell wall of Sitka spruce in relation to hydrostatic stress, *P* (reproduced by permission of Springer-Verlag from "Wood–Water Relations," C. Skaar, 1988, p. 78, Figure 2.26).

expressed as a ratio of adsorption (A) to desorption (D) moisture content for the same relative humidity. The A/D ratio ranges from 0.785 to 0.844 (mean $A/D = 0.812 \pm 0.023$) for softwoods and 0.79 to 0.849 (mean $A/D = 0.828 \pm 0.018$) for hardwoods. Sorption hysteresis decreases with increasing temperature and disappears at temperatures about 75 °C (Skaar 1988).

The immediate effect of higher temperatures on the sorption isotherm is to decrease M_E for the same relative humidity (Fig. 2). M_F is also reduced as the temperature increases (Siau 1984). This effect is reversible. The long-term effect of subjecting wood to higher temperatures is to reduce permanently the hygroscopicity, especially at temperatures above 100 °C. This effect is not reversible when wood is brought back to room temperature. The mechanical properties are also reduced by long-term exposure to high temperatures (Stamm 1964).

The presence of extractives in wood depresses the isotherm at higher relative humidity, typically above 60%, resulting in lower M_F. Isotherms of extracted and unextracted wood coincide very closely at relative humidity below about 60% (Wangaard and Granados 1967).

Mechanical stress affects the sorption isotherm. At constant relative humidity, a compressive stress reduces the sorption M_E and tensile stress increases it. This is known as the Barkas effect (Skaar 1988) and is shown in Fig. 3.

Not all wood is dried to a zero moisture content before use. If a green piece of wood is dried to a moisture content of 12% (humidity approximately 50%) and then subjected to a higher humidity of 85%, the new M_E would follow the general trend of its adsorption curve. Cycling the wood through two humidities at the same temperature will result in a closed ("hysteresis") loop between the desorption and adsorption isotherms (Skaar 1988).

4. *Thermodynamics of Moisture Sorption*

The energy levels of the three forms of water in wood are shown in Fig. 4. Water vapor has the highest level, followed by free water in the cell lumen and voids, which has the same energy as ordinary liquid water. Bound water in the cell walls is in the lowest energy state, increasing with increasing moisture content to M_F. Above M_F, bound water has the same energy level as free water. At low moisture contents, the bound water energy is lower than that of ice. The heat of vaporization of liquid water, E_o (kJ mol^{-1}), is the energy required to evaporate liquid water and E_v is the energy required to evaporate bound water,

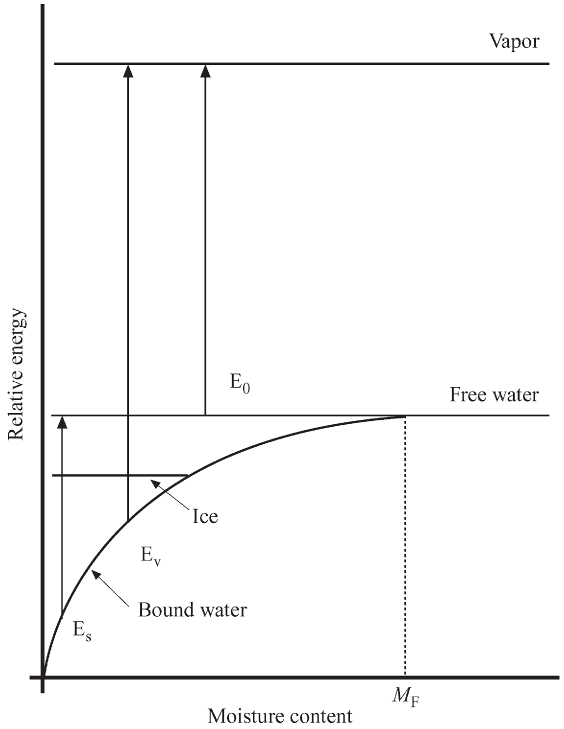


Figure 4
Relative energy levels of the different forms of water in wood as functions of moisture content.

called heat of vaporization of bound water. The difference between these heats of vaporization is called the differential heat of sorption, i.e., $E_v - E_o = E_s$. It is a measure of the excess binding energy of the water molecules to the wood substrate above that between water molecules in the liquid state (Skaar 1988). The differential heat of sorption is strongly dependent on the moisture content of wood, ranging from 0 kJ mol^{-1} at M_F to 22.8 kJ mol^{-1} near 0% moisture content, and has an exponential behavior (Fig. 5). An approximate empirical equation in the hygroscopic range is (Skaar 1988):

$$E_s = \exp(a - bm) \quad (3)$$

where a and b are constants.

Addition of water to oven-dry wood results in an exothermic reaction and the heat generated is defined as the heat of wetting, W, having units of energy per oven-dry mass of wood. At 0% moisture content, the heat of wetting is the "total heat of wetting", W_o, and at higher moisture contents, the difference between W_o and W is defined as the "integral heat of sorption." Both of these terms are related to the differential heat of sorption; to keep units the same, Q_s is

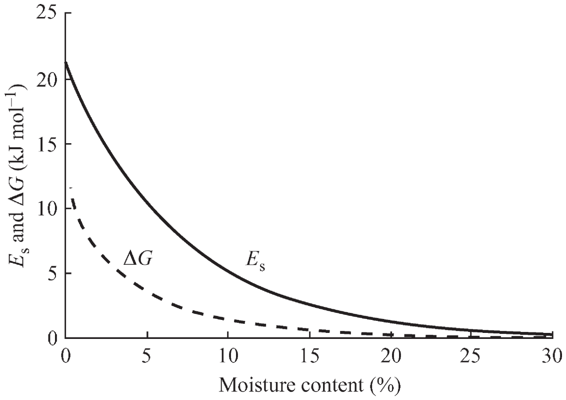


Figure 5
Typical curves of ΔG and E_s as functions of moisture content.

used on a mass basis for differential heat of sorption, equal to E_s divided by the molecular weight of water (0.018 kg mol^{-1}). The relationship of E_s (and Q_s) and W is given by:

$$\frac{E_s}{0.018} = Q_s = \frac{-\mathrm{d}W}{\mathrm{d}m} \quad (4)$$

If Q_s is known then W can be evaluated using Eqn. (3):

$$W = \frac{Q_s}{b} = \left(\frac{1}{b}\right)\exp(a - bm) \quad (5)$$

The total heat of wetting, W_o, is the value when $m = 0$.

The free-energy change, ΔG (kJ mol^{-1}), associated with moisture sorption can be calculated from:

$$\Delta G = RT \ln h \quad (6)$$

where R is the molar gas constant (kJ mol^{-1} K^{-1}) and T is the temperature (K). Since h is less than unity, ΔG is negative, the absolute magnitude decreasing with increasing h (Fig. 5). The corresponding decrease in entropy, ΔS, associated with sorption is given by:

$$\Delta S = \frac{(E_s - \Delta G)}{T} \quad (7)$$

5. *Theories of Moisture Sorption*

5.1 *Unimolecular Sorption*

Langmuir (1918) was the first to develop a sorption theory for localized unimolecular layer adsorption on an energetically homogeneous surface which was then generalized to describe adsorption on real energetically heterogeneous surfaces. Langmuir's equation is:

$$m = \frac{m_\mathrm{m} ch}{1 + ch} \tag{8}$$

where m_m is the moisture content fraction corresponding to the unimolecular layer and c is a constant related to the energy of sorption Theoretical development and extension of the Langmuir equation considered three concepts: attractive interactions between adsorbed molecules; nonelectrolyte solutions; and analogies to bulk systems, where the absorbed gas was considered as a two-dimensional gas exerting a spreading pressure (Rudzinski and Everett 1992). The primary concepts applied to wood sorption are formation of unimolecular and multimolecular layers, mixture of wood water and polymer solution, and formation of water clusters. Venkateswaran (1970) presents a review of mathematical models developed for moisture sorption on cellulosic materials and Skaar (1988) details sorption theories for wood. Most of the models were all derived for adsorption characteristics; however, they are used for describing the desorption curve.

5.2 Multilayer Theories

The multilayer theories are both based on a "multilayering" concept. Figure 6 shows a schematic view of the formation of layers during sorption. The theories propose a unimolecular layer tightly bound to the hydroxyl groups of the wood constituents, and secondary layers which are thermodynamically the same, but different from the unimolecular layer. Brunauer *et al.* (1938) developed the BET theory which assumes a unimolecular layer tightly bonded to the wood substrate, and secondary and higher layers having similar characteristics as liquid water. Dent (1977) modified the BET theory to consider the secondary and higher layers to be thermodynamically the same, but different from that of liquid water. The Dent model has been shown to be a more general model with the BET and Langmuir models being limiting cases. The BET equation is given by:

$$\frac{m}{m_\mathrm{m}} = \frac{ch}{1-h}\left(\frac{1-(n+1)h^n + nh^{(n+1)}}{1+(c-1)h - ch^{(n+1)}}\right) \tag{9}$$

If $n = 1$, it reduces to the Langmuir model. The Dent theory introduces a term to account for a different sorption energy associated with secondary sorbed water molecules. According to this theory

$$\frac{m}{m_\mathrm{m}} = \frac{b_1 h}{(1-b_2 h)(1-b_2 h + b_1 h)} \tag{10}$$

where $b_1 = S_1/(S_0 h)$ and $b_2 = S_i/(S_{(i-1)} h)$, $i \neq 1$. S_i is the area covered by water in layer i and S_0 is the total available sorption area. The term b_2 is a constant that relates the statistical thermodynamic partition function of the secondary sorbed molecules to that of an external condensed liquid in equilibrium with the saturated vapor. In the BET theory, it was assumed that this constant was equal to unity, since the energies of the secondary molecules were the same as liquid water. Term b_1 is defined as the ratio of the partition function of the primary molecules to the external condensed liquid in the saturated vapor. Dent considered the secondary sorbed water to be different from liquid water as well as different from the primary sorbed water. A key assumption is that the partition functions of the secondary sorbed water molecules are the same for all layers.

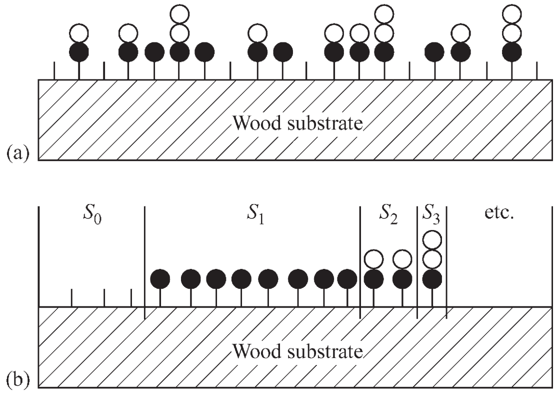


Figure 6
(a) Schematic diagram showing multimolecular layering. The vertical lines represent sorption sites, the filled circles represent primary molecules, and the open circles represent secondary molecules. (b) As in (a) but rearranged in ascending order of site occupancy (reproduced by permission of Springer-Verlag from Harley *et al.*, 1992 Cluster theory for water sorption in wood. *Wood Science and Technology* **26**, 83–99).

Le and Ly (1992) developed a multilayer theory (denoted as the L–L model) similar to the BET theory. The L–L model closely follows the development of the BET model; however, the former has different underlying assumptions. The BET model assumes that the adsorption site area, S_i, is covered by i molecular layers, whereas the L–L model assumes that S_i is the area occupied by the ith layer of the adsorbed molecule. Their model presents a possible way to determine the fractions of primary and secondary sites available during the adsorption process. Based on these assumptions, the following model was developed:

$$m = \frac{m_\mathrm{m} ch}{1+ch}\left(\frac{1-\alpha h^n - (1-\alpha)h^p}{1-h}\right) \tag{11}$$

The fraction of the adsorption sites that contain up to n layers is α and the rest of the sites have up to $p > n$ layers.

5.3 Solution Theory

Hailwood and Horrobin (1946) developed a model for water sorption by assuming that the process is governed by the formation of a solid solution of water in the polymer and the formation of hydrates between water and definite units of the polymer molecule. The mixture of polymer, polymer hydrates, and dissolved water forms a single solid phase. The two-component system, water and polymer, exists in two phases, and the dissolved water molecules in the solid phase are assumed to be mobile. A one-hydrate model represents the experimental sorption data quite well, given by:

$$M_{\mathrm{p}}\frac{m}{18}=\frac{\alpha h}{1-\alpha h}+\frac{\alpha\beta h}{1+\alpha\beta h} \qquad (12)$$

where $\alpha = Ka/h$, $\beta = K_1$, M_{p} is the molecular weight of polymer, K is an appropriate equilibrium constant between the different phases of water, and a is the activity of the water (essentially the relative vapor pressure). The first term on the right-hand side of Eqn. (12) represents the fraction of free dissolved water and the second term the various fractions of water present as monohydrate. The coefficients K and K_1 have been determined as functions of temperature for purposes of obtaining a generalized sorption isotherm at different temperatures (Skaar 1988).

5.4 Cluster Theory

Hartley and Avramidis (1993) used statistical thermodynamics to describe the formation of water clusters during adsorption. Although there is no model developed to form the basis of adsorption, it provided insight into the mechanism of adsorbing water in wood. During the adsorption process, as the moisture content increases, water molecules are randomly adsorbed, then “bridges” and small clusters of two or three water molecules are formed at $h > 0.35$. Finally, larger clusters of four to six molecules form at higher relative vapor pressures, $h > 0.9$. The process is not temperature dependent; however, changes of sizes at relative vapor pressures occur at lower h as temperature increases.

See also: Wood, Constituents of*; Wood: Density#; Wood: Diffusion and Permeability#; Wood: Dimensional Change from Moisture#

Bibliography

American Society for Testing and Materials 1998 Standard test methods for direct moisture content measurement of wood and wood-base materials. Designation D 4442-92 (Re-approved 1997). ASTM, Philadelphia, PA

Brunauer S, Emmett P H, Teller E 1938 Adsorption of gases in multimolecular layers. *J. Am. Chem. Soc.* **60**, 309–19

Dent R W 1977 A multilayer theory for gas sorption: I. Sorption of a single gas. *Textile Res. J.* **47**, 145–52

Hailwood A J, Horrobin S 1946 Absorption of water by polymers: analysis in terms of a single model. *Trans. Faraday Soc.* **42B**, 84–92

Hartley I D, Avramidis S 1993 Analysis of the wood sorption isotherm using clustering theory. *Holzforschung* **47**, 163–7

Hartley I D, Kamke F A, Peemoeller H 1992 Cluster theory for water sorption in wood. *Wood Sci. Technol.* **26**, 83–99

Haygreen J G, Bowyer J L 1996 *Forest Products and Wood Science: An Introduction.* Iowa State University Press, Ames, IA

Langmuir I 1918 The adsorption of gases on plain surfaces of glass, mica, and platinum. *J. Am. Chem. Soc.* **40**, 1361–405

Le C V, Ly N G 1992 Multilayer adsorption of moisture in wool and its application in fabric steaming. *Textile Res. J.* **62** (11), 648–57

Panshin A J, de Zeeuw C 1980 *Textbook of Wood Technology.* McGraw-Hill, New York

Rudzinski W, Everett D H 1992 *Adsorption of Gases on Heterogeneous Surfaces.* Academic Press, London

Siau J F 1984 *Transport Processes in Wood.* Springer, Berlin

Skaar C 1988 *Wood–Water Relations.* Springer, Berlin

Stamm A J 1964 *Wood and Cellulose Science.* Ronald Press, New York

Venkateswaran A 1970 Sorption of aqueous and non-aqueous media by wood and cellulose. *Chem. Rev.* **70**, 619–37

Wangaard F F, Granados L A 1967 The effect of extractive on water-vapor sorption by wood. *Wood Sci. Technol.* **1**, 253–77

I. Hartley
Forintek Canada Corp., Vancouver, British Columbia, Canada

Wood: Nonstructural Panel Processes

Nonstructural panels include particleboard and fiberboard, with the difference being mainly in the geometry of the wood elements that comprise them. Fiberboard can be further categorized as hardboard, medium density fiberboard (MDF), and insulation board. Density is the differentiating factor between the fiberboard types, as well as some important processing steps. Hardboard has a density of 900 kg m^{-3} or more, MDF between 500 and 900 kg m^{-3}, and insulation board between 150 and 500 kg m^{-3}. The categorizing of panel types is not always clear. Composite panel manufacturing allows for many combinations of particles and fibers, a broad range of density, with a variety of adhesive types. Nonstructural panels can be engineered to meet very specific performance criteria (see *Wood: Nonstructural Panels**).

The term panel usually refers to a flat panel. However, any of the nonstructural panels can be molded into a three-dimensional configuration. The following discussion is limited to flat-pressed, wood-based panels, although other plant-based materials may be used. Wheat straw is used in North America to produce particleboard, as well as bagasse, bamboo, flax, and others in many parts of the world. The basic manufacturing steps are the same.

1. *Raw Materials*

1.1 Wood

Sources for wood used in the manufacture of nonstructural panels vary depending on availability and desired panel properties. Some roundwood is used in order to improve control over the particle or fiber geometry; however, roundwood is comparatively expensive. The most common sources of wood are residues from lumber and plywood manufacturing processes, including planer shavings, sawdust, trim, and other mill residues. Pulp chips are used frequently in the manufacture of fiberboard products. Both hardwood and softwood species can be used. The adhesive and manufacturing parameters can be adjusted to accommodate many wood species.

1.2 Adhesive

Urea–formaldehyde (UF) is the most common adhesive used in the manufacture of particleboard and MDF. This adhesive is nearly colorless, provides a strong bond, has a fast cure time, and is relatively inexpensive. However, UF adhesive is not stable in wet or high-humidity environments, which limits its use to interior and nonstructural applications. In some cases melamine–formaldehyde (MF) is used for better moisture resistance, either by itself or mixed with UF adhesive. Some phenol–formaldehyde (PF) and methyl diisocyanate (MDI) adhesives are used in the manufacture of particleboard and MDF. These adhesives are more expensive than UF but are resistant to moisture. Panels using these adhesives are intended for applications where moisture intrusion is a possibility, and comprise a very small part of the nonstructural panel market (see Wood: Adhesives[#]).

Hardboard is generally produced using a small amount of PF adhesive (1–2% of the panel weight), although it may be produced using no adhesive. High temperature and pressure during hot pressing promotes ligneous bonding. Hydrogen bonds, and perhaps covalent bonds, are initiated in the hot-pressing step as the thermoplastic lignin component of the wood substance flows, allowing intimate contact between the fibers. It has been proposed that some of the hemicellulose polymers degrade to yield reactive sites where covalent bonds could form. The addition of PF adhesive to hardboard improves its strength, stiffness, and moisture resistance.

Bonding between the fibers in insulation board is due partially to hydrogen bonds. Usually starch or asphalt is added to the fibers to provide some adhesive bonding. Since no hot-pressing step is used in the manufacture of insulation board, and the panel density is low, the bond strength is very weak.

1.3 Wax

A petroleum-based wax is typically added to the particles or fibers during the manufacture of nonstructural panels. The purpose of the wax is to impart some resistance to the penetration of liquid water into the panel. The wax provides no resistance to the penetration of water vapor. The wax may be added in a molten state, or more commonly emulsified and added as a liquid at room temperature. The wax content is quite small, perhaps 0.5–1.0% based on the weight of the panel. Other sizing agents are often used in hardboard and insulation board, including rosin and asphalt.

1.4 Other Additives

Other materials are sometimes added to the particles or fibers depending on the end use. Fire-retardant chemicals and preservatives may be added prior to mat formation. Some hardboard is tempered, which involves the addition of oil before or after (usually) hot pressing. Linseed, tung, soybean, and tall oils are used typically for tempering. In the case of UF adhesive, additional urea may be added as a scavenger of excess formaldehyde during the manufacturing process.

2. *Particleboard Processing*

2.1 Particle Preparation

Depending on the starting raw material, the preparation of particles suitable for particleboard may involve a variety of machines and processing steps (Fig. 1). In general, these machines employ knives, hammers, or grinding disks to break down the wood into small pieces of the desired size and shape. Small particles are desired for the surfaces of particleboard to impart a smooth and uniform texture. Particle sizes are usually defined in terms of a screen classification, since the particles have an irregular shape. Surface particle sizes may range from 0.5 mm down to individual fibers of approximately 50 μm. Particles to be used in the core of particleboard are thicker and longer, and usually much more variable in size and shape than surface particles. Core particles are long and slender, with length of perhaps 1–10 mm and thickness of 0.5–1 mm.

Roundwood (logs) must have the bark removed in a debarker prior to particle preparation. The presence of bark in particleboard generally reduces the product quality. The logs are then fed directly to a long-log flaker to produce relatively slender particles, or cut into short bolts to be fed to a chipper. Larger

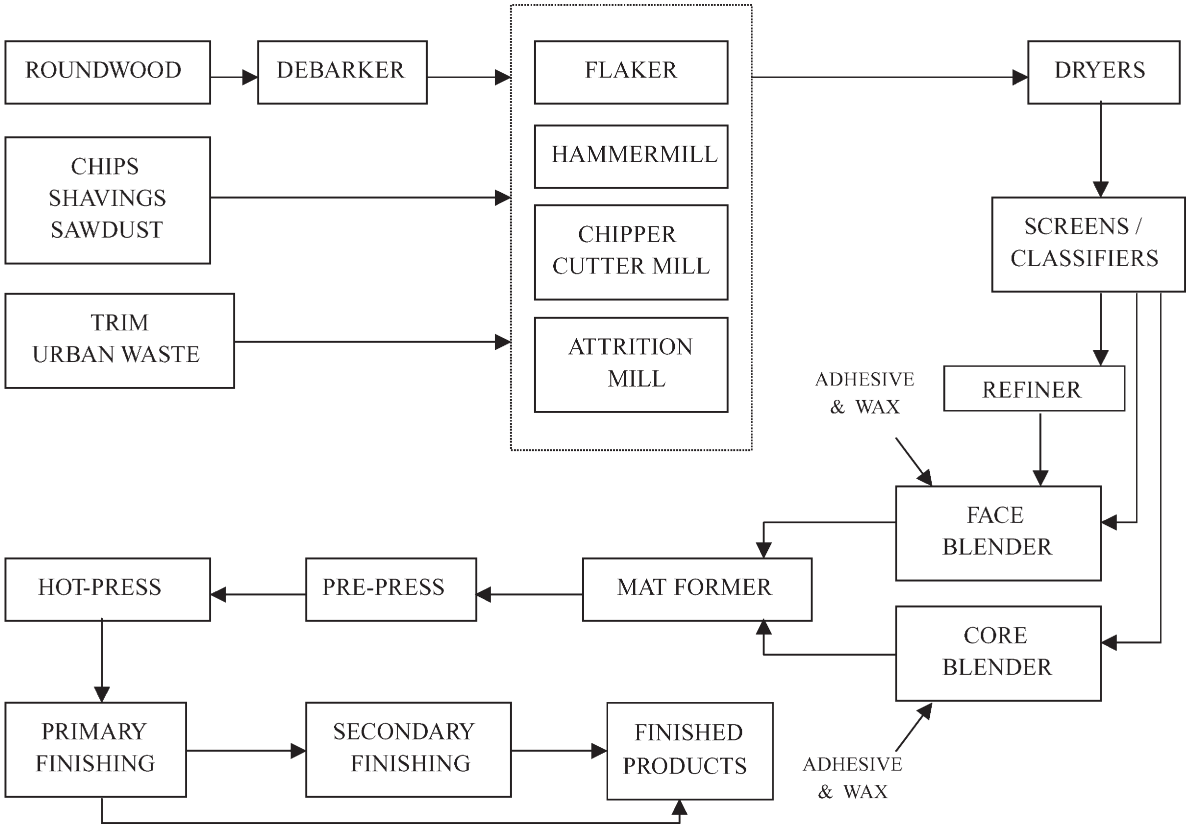

ROUNDWOOD
DEBARKER
CHIPS
SHAVINGS
SAWDUST
TRIM
URBAN WASTE
FLAKER
HAMMERMILL
CHIPPER
CUTTER MILL
ATTRITION
MILL
DRYERS
SCREENS /
CLASSIFIERS
REFINER
ADHESIVE
& WAX
FACE
BLENDER
CORE
BLENDER
ADHESIVE
& WAX
HOT-PRESS
PRE-PRESS
MAT FORMER
PRIMARY
FINISHING
SECONDARY
FINISHING
FINISHED
PRODUCTS

Figure 1
Schematic diagram of particleboard manufacture.

sawmill residues, such as end- and edge-trim, may also be fed to a chipper to produce a wood chip. Such a chip must be further reduced to particles or fibers.

Smaller wood residues, such as plywood trim, chips, planer shavings, and sawdust, may be fed directly to a hammer mill, flaker, or attrition mill. Hammermills and impact mills are crude breakdown devices that literally beat the wood into smaller pieces. Flakers, of which there are several designs, use knives with precise set points to produce a uniform particle geometry. Attrition mills, also called refiners or "fiberizers," employ a shearing action to tear wood into bundles of fibers. In an attrition mill the wood to be processed is forced between two opposing disks, with one or both disks revolving at a high velocity. The disks are machined with ribs and channels to impart a shearing action when the wood passes between them. Attrition mills usually are operated in conjunction with a pressure vessel, which preheats the wood in steam for the purpose of softening the fibers prior to refining. Not all particleboard processing uses attrition mills. However, the fine wood-fiber bundles are used for the surface layers of particleboard.

Particles are segregated typically by size using a variety of mechanical screen systems. Usually face (smaller) particles and core (larger) particles are separated. Also the very small particles, called fines, are removed and used for fuel, and particles that are too large for the core are further refined and placed back into the process. Screening may occur prior to drying, or more likely after drying.

2.2 Particle Drying

Moisture content of the wood particles is a critical factor in the manufacture of particleboard. Moisture content influences adhesive bonding, heat transfer in the hot-pressing step, and the development of density variations in the panel. The most common method of wood particle drying is the rotary dryer. Rotary dryers consist of a large rotating drum, with its rotating axis in the horizontal position. In the drum the wet wood particles are in direct contact with the hot drying gases. The drying gases are a mixture of combustion gases, usually from a wood-fired burner, and ambient air. The drying gases serve two purposes in this continuous drying process: heat transfer to the particles and pneumatic conveyance of the particles through the drum. The slow rotation of the drum promotes mixing of the particles with the drying gases.

2.3 Particle Blending

Blending of particles with liquid adhesive is typically performed in a short-retention-time blender. These devices consist of a fixed, horizontal cylinder of less than 1m in diameter and 3–4 m in length. As the particles enter the cylinder they are agitated with a high-speed rotor and a series of paddles, which move the particles through the blender. Liquid adhesive is applied using spray nozzles near the particle inlet. Much of the adhesive distribution is accomplished by the tumbling action and rubbing together of the particles. Wax and other additions may also be applied in the same blender as the adhesive, either with separate spray nozzles, or mixed with the adhesive prior to the blender.

2.4 Particle-mat Forming

After the blending process, the particles, now referred to as furnish, must be formed into a mat. Forming is accomplished with a system called a forming machine, which consists of a series of bins, spreaders, and a continuous conveyor that immediately precedes the hot press. Core and face furnish are fed to separate bins, which in turn supply spreader heads that are positioned above the conveyor. The spreader heads distribute the furnish onto the conveyor in layers.

A three-layer particleboard manufacturing process would have two surface-layer bins, each with a spreader head, and one core-layer bin, with a spreader head. Five-layer, or more, particleboard would have more bins and spreader heads. The forming machine controls the amount of furnish deposited for each layer. Particleboard properties can be manipulated by changing the composition and amount of each layer. Adjusting the adhesive content and particle geometry of the furnish in each layer, along with the adjustments at the forming machine, permit the particleboard properties to be engineered for very specific applications.

2.5 Particleboard Pressing

The purpose of the hot press is to consolidate the loose mat of furnish into a panel. The hot press provides the mechanical force necessary to bring the particles together and the thermal energy needed to cure the adhesive. The amount of press time needed depends on the thickness of the panel, temperature of the press platens, and the type of adhesive. Three to five minutes is typical, for a thickness of 12–18 mm, when using a platen temperature of 130–160 °C. Both continuous, belt-type hot presses and multiopening, platen-type hot presses are used.

2.6 Postpressing Operations

After hot pressing UF-bonded particleboard is quickly cooled to prevent thermal breakdown of the adhesive and allow off-gassing of formaldehyde. PF-, MF-, and MDI-bonded particleboard does not require cooling. The panels are then edge-trimmed and cut to size. The panels are visually graded using industry standards, and a testing procedure is followed rigidly to ensure product quality. Testing includes strength, stiffness, moisture resistance, and dimensional stability.

A variety of secondary processing operations, such as sanding, milling, surface lamination, direct printing, and embossing, also may be performed.

3. MDF Processing

There are many similarities between the manufacture of MDF and particleboard. Indeed, these products are used for many of the same products. The major process differences occur in the fiber preparation, drying, and blending operations (Fig. 2). See Sect. 2 for a description of the pressing and postpressing operations.

3.1 Fiber Preparation

MDF fiber is produced from chips and shavings using a pressurized refining system. This is essentially a thermomechanical pulping process. The fiber is actually a collection of fiber bundles and individual wood cells. Hardwood and softwood species will produce substantially different MDF fiber due to the anatomical differences in regard to cell type and density.

3.2 Fiber Blending

Adhesive and wax are added to the fiber either in the refiner or in the blowline. Blowline blending is the most common, and consists of an injection nozzle that is positioned inside the pneumatic duct between the refiner and the dryer. The distribution of the adhesive and the wax is accomplished to some degree by the atomization of the adhesive, but also due to the turbulent flow and mixing that occurs in the blowline after adhesive application. Some MDF plants may use a short-retention-time blender, such as employed for particle blending, instead of the blowline blending method.

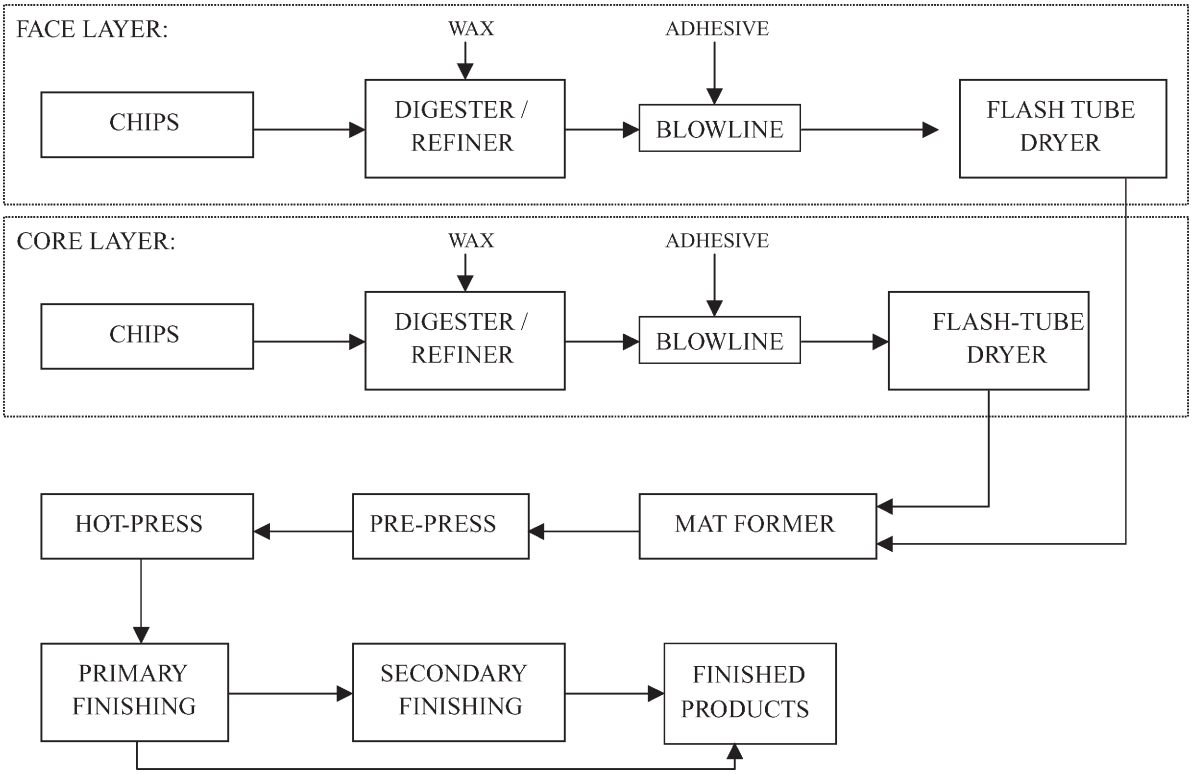


Figure 2
Schematic diagram of medium density fiberboard manufacture.

3.3 Fiber Drying

Fiber dryers, sometimes called flash-tube or tube dryers, are pneumatic drying systems. Other than the blowers, there are no moving parts. The wet fibers are entrained by the hot drying gases at the inlet of the dryer and pushed through a long tube, approximately 50–100 m in length and 1m in diameter. At the end of the dryer the dry fibers are separated from the high velocity air using a cyclone separator. Retention time in the dryer is short enough to prevent cure of the adhesive.

3.4 MDF Forming

MDF mats are formed in a manner similar to particleboard. However, the spreader heads are replaced with felting heads, which break up clumps of fibers before they are deposited onto the conveyor. The mat formation is assisted with the application of a slight vacuum beneath the conveyor. In this case the conveyor is a porous screen, onto which the mat is formed. The vacuum allows the fibers to be filtered out of the air-stream and uniformly deposited on the conveyor. MDF mats are typically formed in three layers.

The remainder of the MDF manufacturing process is similar to particleboard manufacture.

4. Hardboard and Insulation Board Processing

The wood fiber used in the manufacture of hardboard and insulation board is produced in a manner similar to MDF fiber. Once the fiber is produced in the pressurized refining process there are three primary techniques for producing hardboard (Fig. 3). The difference is in the method of mat formation and the moisture content at the time of pressing. These methods are termed wet–wet, wet–dry, and dry–dry. Another difference between the hardboard processes is whether adhesive is added, which is not required for wet–wet process hardboard.

In the wet–wet method the mat is formed wet and consolidated in the hot press when wet. Starting with the wet fiber from the refiner, more water is added to produce a slurry of water and wood fiber. Although some adhesive may be added at this point, none is needed. A mat is then formed on a Fourdrinier machine. This procedure is similar to the manufacture of paper (see *Papermaking**). Some of the water is mechanically squeezed out of the mat prior to hot

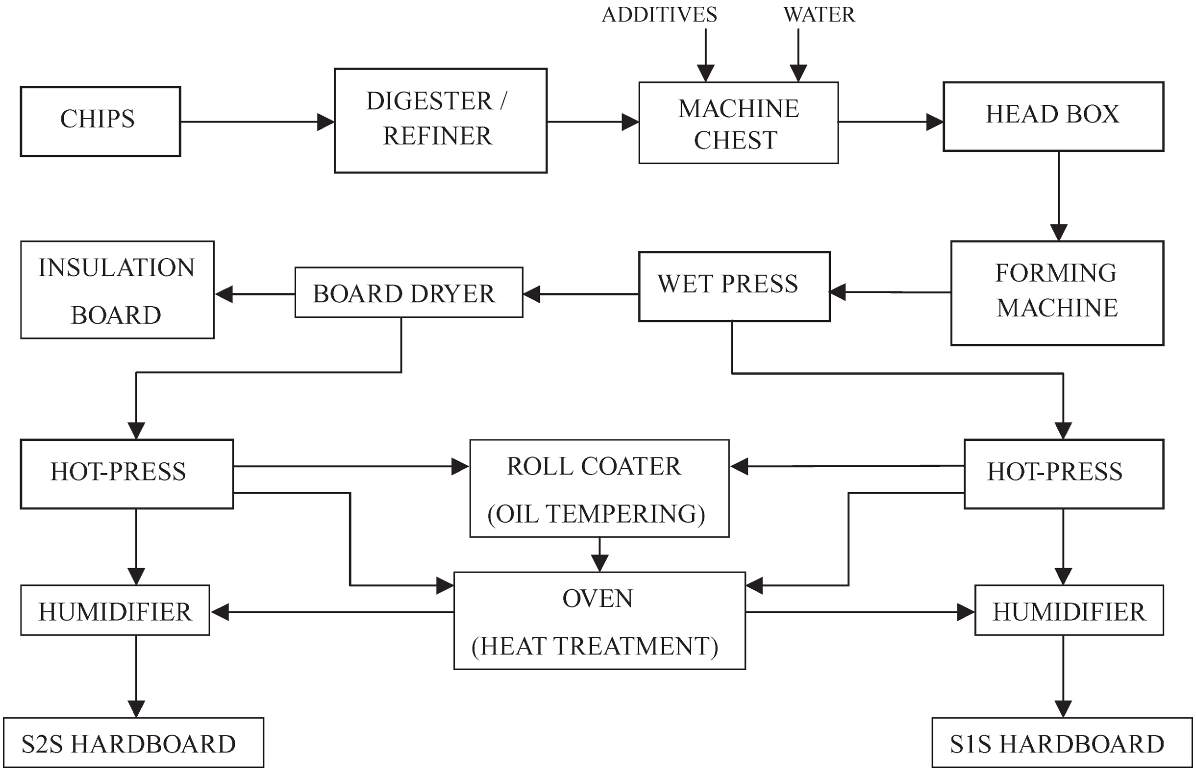


Figure 3
Schematic diagram of hardboard and insulation board manufacture.

pressing, but the moisture content remains above 100% (dry basis). High temperature and high compaction pressure in the hot press results in a high-density panel. The bonding between the fibers is primarily the result of lignin flow and hydrogen bonds. Since the mat is wet at the start of hot pressing, it must be pressed on a wire mesh to allow the escape of steam. This causes an impression of the wire mesh on one side of the resulting panel, or what is called a smooth-one-side (S1S) panel.

The wet–dry process consists of the wet-forming step, but undergoes a mat-drying step prior to hot pressing. With the removal of the water before hot pressing the mat may be pressed on smooth, metal platens. This will produce a panel that is smooth-two-sides (S2S). As with the wet–wet process, adhesive is not needed due to the high temperature and high compaction pressure in the hot press, which promote natural bonding.

The dry–dry hardboard manufacturing process is similar to the MDF manufacturing process. The primary difference is that dry–dry hardboard is pressed to a greater density than MDF. In addition, a PF adhesive is used in dry–dry hardboard, whereas a UF adhesive is usually used in MDF.

Numerous secondary manufacturing processes may be employed with hardboard. Oil tempering and/or heat treatments are often used to improve moisture resistance. Overlay materials are commonly bonded to hardboard for such applications as paneling and furniture parts. Embossing and polymer coatings may be added for exterior siding.

The manufacture of insulation board follows the same procedure as is used for wet–dry hardboard, with the omission of hot pressing. Starch or asphalt is added to the furnish to enhance fiber bonding in this low-density product.

5. *Summary*

This article summarizes the basic processing steps in the manufacture of the most common types of nonstructural, wood-based panels. There are numerous panel configurations that are manufactured, as well as countless panel products that may be devised in the future. There is much attention paid to the blending of wood with nonwood materials, both natural plant fibers and synthetics, to produce new panel products. The nonstructural panel products are produced in most parts of the world, and there are very few regional differences in the manufacturing technology. The classification of these panels is becoming less distinct, with greater emphasis on product performance than the method of manufacture or the raw materials.

Bibliography

Haygreen J G, Bowyer J L 1996 *Forest Products and Wood Science: An Introduction*. Iowa State University Press, Ames, IA

Maloney T M 1977 *Modern Particleboard and Dry-process Fiberboard Manufacturing*. Miller Freeman, San Francisco

Plackett D V, Dunningham E A 1992 *Proc. Pacific Rim Bio-based Composites Symp*. Forest Research Institute Bulletin No. 177

Suchland O, Woodson G E 1987 *Fiberboard Manufacturing Practices in the United States*, Agriculture Handbook No. 640. US Department of Agriculture, Washington, DC

F. A. Kamke
Virginia Polytechnic Institute and State University, Blacksburg, Virginia, USA

Wood: Nonstructural Panels

Wood-based composite panels are constructed from wood and synthetic binder for a variety of applications. Most of these applications require mechanical performance but only those in building construction (i.e., sheathing, sub-floor, web-stock, etc.) are classified as structural. Often, nonstructural applications require a significant level of mechanical performance to function correctly. However, that performance is not regulated beyond product performance standards and poses low risk in the event of failure. Typical applications for nonstructural panels include furniture, case goods, interior doors, decorative paneling, and moldings.

These materials are developed to mimic the performance of wood with flat plate geometry. Typically, the panels are composed of wood in various forms that are bonded with thermosetting polymeric resins. Various manufacturing processes that center on flat pressing are employed and are discussed elsewhere. This article focuses on the definition, description, and use of lignocellulosic panels for various nonstructural uses.

1. Types of Panels

Both the production process and wood component form define a nonstructural wood composite. The majority of the composites are manufactured via a dry flat press process. Alternative systems such as wet processing or extrusion, although technically applicable, are not the preferred methods of manufacture. Of the dry process nonstructural wood composites, four dominate: particleboard, hardboard, fiberboard (medium density), and decorative plywood (Fig. 1). Each of these composites can be found named and defined differently throughout the world. Below are definitions as offered by the American Society for Testing and Materials (ASTM) Standard D 1554 "Standard Definitions of Terms Relating to Wood-Base Fiber and Particle Panel Materials," the National Particleboard Association Standard NPA 4-73 "Standard for Medium Density Fiberboard," and the Canadian Standards Association Standard O115-M1982 "Hardwood and Decorative Plywood."

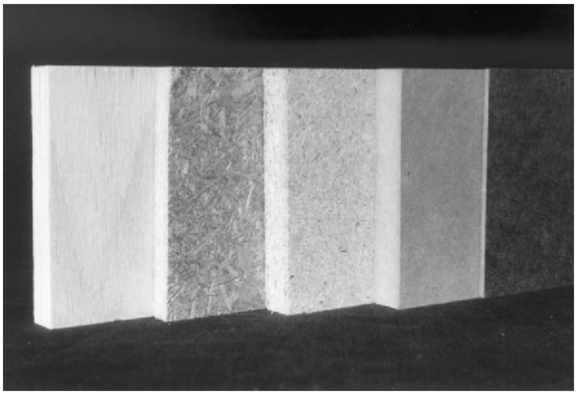


Figure 1
Various nonstructural wood–composite panels (left to right decorative plywood, flakeboard, particleboard, medium density fiberboard, and hardboard).

1.1 Particleboard

A generic term for a panel manufactured from lignocellulosic materials (usually wood) primarily in the form of discrete pieces or particles, as distinguished from fibers, combined with a synthetic resin or other suitable binder, and bonded together under heat and pressure in a hot-press by a process in which the entire interparticle bond is created by the added binder, and to which other materials may have been added during manufacture to improve certain properties (ASTM D 1554).

1.2 Hardboard

A generic term for a panel manufactured primarily from interfelted lignocellulosic fibers (usually wood), consolidated under heat and pressure in a hot-press to a density of 31 lb ft^{-3} (500 kg m^{-3}) or greater, and to which other materials may have been added during manufacture to improve certain properties (ASTM D 1554).

1.3 Medium-density Fiberboard (MDF)

A dry formed panel product manufactured from lignocellulosic fibers combined with a synthetic resin or other suitable binder. The panels are compressed to a density of 31–50 lb ft^{-3} (500–800 kg m^{-3}) in a hot press by a process in which substantially the entire interfiber bond is created by the added binder. Other materials may have been added during manufacture to improve certain properties (NPA 4-73).

1.4 Decorative Plywood

A hardwood or softwood assembly made of three or more layers of wood veneers jointed together with an adhesive, of which one or both of the outer plies comprise veneers that have appearance as the primary consideration (CSA O115).

2. Raw Materials

2.1 Wood—Agricultural Residues

The manufacture of nonstructural wood composite panels is a global industry. As such the species of wood/agricultural fiber used for their production is dependent on available resources and is diverse. For composites such as particleboard, hardboard, and medium-density fiberboard, softwood (gymnosperm) species, namely pine, spruce, and Douglas fir dominate (Composite Panel Association 1999–2000). However, various other softwood/hardwood species, e.g., beech and birch, as well as agricultural crop residues are also regularly used (Composite Panel Association 1999–2000). Generally, panels will not be manufactured from one species alone. Rather a carefully controlled mixture of species will be metered into the production process. An example of the species mix that could be common to a European particleboard manufacturing facility is listed in Table 1.

Table 1
Possible species mix of a European particleboard manufacturing facility.

Wood species	Species proportion %
Sitka and Norway spruce	30
Lodgepole, Scots, and Corsican pine	25
Radiata pine	–
Southern pine	–
European and Japanese larch	15
Mixed conifer	20
Mixed broadleaf	10

Traditionally, the nonveneer wood composites industry has operated on residues from the forest products industry (Maloney 1997). Material of various forms: roundwood, sawmill residues and even recycled wood fiber can be used for panel manufacture. In countries other than the USA, the use of roundwood is common. However, because of their low cost and abundant supply, sawmill residues are the dominant raw material in the USA (Maloney 1997). As technology has advanced, and fears have grown over deforestation, crop stubble burning, and other environmental issues, the use of agricultural residues as an alternative feedstock for nonveneer composite panels has also increased. Currently, crop residues such as wheat straw, bagasse (sugar extracted sugar cane), or ryegrass straw are all being used in the production of nonstructural wood composites. Irrespective of the form or species of the raw material to be used, various stages of furnish preparation are required before panel production is possible.

The decorative plywood industry is not able to make use of material gathered from sawmills, recycling schemes, or agricultural residue. Veneer peeled from solid timber is the mainstay of the industry. However, the species of this timber is still very much dependent on the resources available to the manufacturer, and will, therefore, vary throughout the world.

2.2 Adhesives

Due to the variety of applications in which nonstructural wood composites are used, the adhesive systems used to bind these products vary. Generally, four types of binder are used: urea formaldehyde (UF), melamine formaldehyde (MF), phenol formaldehyde (PF), and polymeric methyl diisocyanate (MDI). Of these, UF is the least expensive and most widely used. Maloney (1997) states that of the world's dry process composition board products, approximately 90% incorporate UF as their binder. The durability of composites made with such an adhesive is generally poor. Therefore, products incorporating UF-bonded composites can not be used in an environment where the risk of water contact is high (Dinwoodie 1981).

MF resin, although more expensive than UF, is also more resistant to moisture. This enables products bonded with MF to be used in situations where occasional product dampness may be a possibility (Maloney 1997). It is increasingly common for UF resins to be fortified with MF, so called MUF resins. Such blends, although more expensive than the UF, exhibit enhanced performance, allowing the finished product to be sold at a higher price.

PF resin offers far greater durability than UF, MF, or MUF systems. Products incorporating PF are deemed appropriate for exterior applications. There are two distinct types of PF resins available: resol or one-stage and novolac or two-stage. Generally, resol PF resins are used for the manufacture of nonstructural wood composites; however, the commercial use of novolac PF resins is not unknown.

MDI resins, although relatively new and expensive, are gaining market share. The durability of such a resin system is particularly good. With the emergence of agricultural residue based products, MDI has proven to be the most appropriate binding option (Heslop 1997). Interest in MDI has also been enhanced by worries over the possible dangers associated with formaldehyde emissions from resins such as UF.

The resins outlined above could all be used for the manufacture of decorative plywood. However, the commonly used resins are PF, UF, MUF, and polyvinyl acetate (Sellers and Haupt 2000). As with nonveneer wood composites, the intended end use of decorative plywood will go a long way in determining the most appropriate resin system.

2.3 Additives

Various additives may be introduced to the composition during manufacture. The most common being moisture retardants, fire retardants, or decay inhibitors. The most frequently used additive is a moisture retardant or sizing agent, usually a wax emulsion. The sizing agent retards the rate of water penetration, safeguarding the composite against accidental wetting, while also aiding secondary gluing operations (Maloney 1997). The addition of a sizing agent also greatly improves composite performance during the standard 24 hour water soak test.

Fire retardants, although less common than moisture retardants, are occasionally added to composite panels. However, their use can also introduce other detrimental effects such as increased composite hygroscopicity (Maloney 1997). Compounds such as zinc borate, as well as inhibiting fire, also have the added benefit of inhibiting fungal decay. Decay inhibitors, although not commonly used in the USA, are popular and even required in other countries, e.g., Germany. Fire retardants, decay inhibitors, and moisture retardants can be added at various stages of panel manufacture. However, it is most likely that each will be added at the same time as the resin.

3. Panel Structure

Nonstructural wood composites generally have a layered construction. Layering is achieved by the variation of the wood components (Fig. 2) and/or the production of what is termed a vertical density profile. Composites such as fiberboard and hardboard are generally homogenous in particle size, i.e., fiber or fiber bundles. For this reason, fiberboard and hardboard panels are generally referred to as being of single layer construction. However, layering of a fiberboard panel with high-density surfaces and a low-density core is common because of the production of a vertical density profile during hot-pressing (Fig. 3).

Composites such as particleboard are more commonly heterogeneous in their particle size, finer particles being laid on the top and bottom surfaces of the panel, and coarser particles being laid within the core region. This type of panel construction is referred to as being three layer (Fig. 4). By itself, this type of construction will lead to a panel with high-density surfaces and a low-density core. However, the layering effect is further compounded by the production of a vertical density profile during hot-pressing.

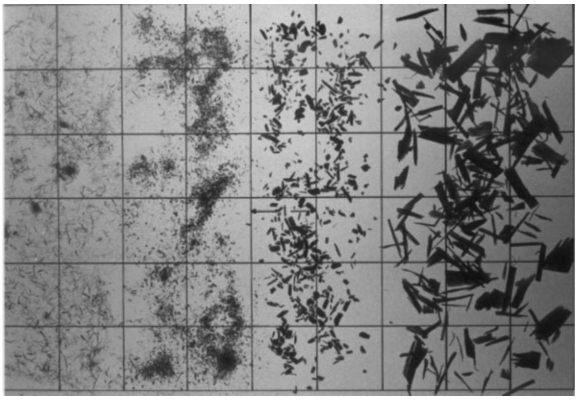


Figure 2
Various wood–composite particle types (left to right fiber, surface layer particles, core layer particles, and flakes).

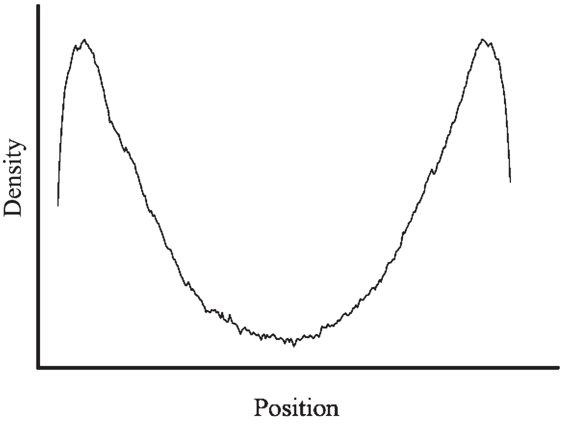


Figure 3
Typical vertical density profile of a hot pressed nonveneer wood composite.

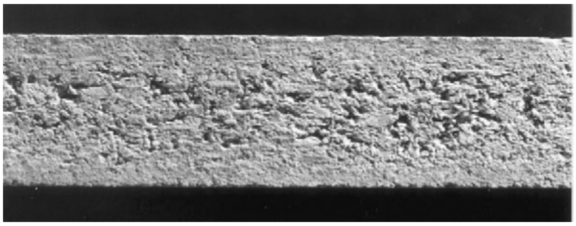


Figure 4
Layered construction of modern day particleboard (panel edge).

The production of a vertical density profile is due to a number of inter-related phenomena. As the hot-press makes contact with the mat, the top and bottom surfaces are consolidated and the resin cured, before the core region has been consolidated to final thickness. It is also common that the surfaces of a mat are of higher moisture content than the core, meaning that these areas are more easily consolidated than the core. Finally, smaller particles (if used for the surface layers) will merge more easily than larger particles, reducing interparticle voids, once again increasing surface density.

By altering the layered construction of a composite panel, its mechanical performance can be significantly altered. In general, those panels of higher surface density and lower core density will tend to exhibit better flexural properties and worse internal bond strength when compared to panels of similar overall density/resin content but more uniform vertical density profile. Variations in resin type and/or loading levels within a panel will also influence panel performance. It is common practice to use different adhesive systems, or application levels for different layers of a panel. A good example of this is the use of resins with faster cure rates in the core of a panel, enabling the various resins throughout a panel to cure as one, minimizing the time the panel spends in the hot-press. It is also common for more resin to be used in areas where fine particles (large specific surface area) are located.

Layering of particle size is also important for secondary processing. A fine, well packed surface layer will lend itself to good surface lamination or painting as telegraphing of large particles will be avoided.

Decorative plywood is also layered, the odd number of wood veneers being stacked with the grain of each being aligned at 90 ° to the next. It is important to note that any panel, regardless of its construction, should be symmetrical through its thickness. If it is not, warping of the panel is inevitable.

Irrespective of panel type, the layering a panel may or may not have prior to hot-pressing, or the vertical density profile desired after pressing, the hot-pressing procedure is uniquely important to panel performance. It is through this process that a composite's wood components are brought into contact, an intermediary layer of resin is allowed to harden or cure, and a

Table 2
Property requirements of particleboard.

Standard	Grade	Modulus of rupture (MPa)	Modulus of elasticity (MPa)	Internal bond (MPa)	Hardness (N)	1 hour thickness swell (%)	Linear expansion max avg. (%)	Face screw holding (N)	Edge screw holding (N)	Formalde-hyde maximum emissions or potential
ANSI A208.1–1993	M-3	16.5	2750	0.55	2225	N/A	$\leqslant 0.35$	1100	1000	0.30 ppm
AS/NZS 1859.1:1997	Standard	16.0	2500	0.40	N/A	$\leqslant 12$	N/A	500	450	0.12 (mg m^{-3})h^{-1}
EN 312–2:1996	Standard	13.0	1600	0.35	N/A	N/A	N/A	N/A	N/A	$\leqslant 30$ mg/100 g

Table 3
Property requirements of medium density fiberboard.

Standard	Grade	Modulus of rupture (MPa)	Modulus of elasticity (MPa)	Internal bond (MPa)	24 hour thickness swell (%)	Linear expansion max avg. (%)	Face screw holding (N)	Edge screw holding (N)	Formaldehyde maximum emissions
ANSI A208.2–1994	Interior MD	24.0	2400	0.60	N/A	N/A	1445	1110	0.30 ppm
AS/NZS 1859.2:1997	Standard	30.0	2800	0.60	$\leqslant 12.0$	$\leqslant 0.4$	1050	850	$0.12\,(\mathrm{mg\ m^{-3}})\mathrm{h}^{-1}$
EN 622–5:1997	Standard (dry conditions)	22.0	2500	0.60	$\leqslant 2.0$	N/A	N/A	N/A	N/A

Table 4
Property requirements of hardboard.

Standard	Grade	Modulus of rupture (MPa)	Internal bond (MPa)	Thickness swell (%)	Water absorption based upon weight (%)	Tensile strength parallel to surface (MPa)	Tensile strength perpendicular to surface (MPa)
ANSI/AHA A135.4–1995	2. Standard	31.0	N/A	$\leqslant 25.0$	$\leqslant 35.0$	15.2	0.62
AS/NZS 1859.4:1997	Standard Type GP	28.0	N/A	$\leqslant 30.0$	N/A	N/A	N/A
EN 622–2:1997	Standard (dry conditions)	30.0	0.50	$\leqslant 30.0$	N/A	N/A	N/A

structure is formed that will be able to withstand the stresses subsequently applied to it.

4. Properties and Standards

Measurement of nonstructural wood–composite performance can take many forms. Composite strength is an obvious example: static bending strength, tensile strength perpendicular to the surface of the panel (generally referred to as internal bond strength), tensile strength parallel to the surface of the panel, compression strength parallel to the surface of the panel, shear strength in the plane of the panel, or edgewise shear strength. Other properties such as: static bending stiffness, fastener holding capacity, water sorption, thickness swell, linear expansion, hardness, abrasion resistance, surface smoothness, and general appearance are also important, and can limit composite use.

Numerous organizations around the world have established and constantly revise standards by which composite quality is determined. Each require various panel properties to be measured in order to characterize performance and several of these properties are common to most standards. Depending upon the overall performance of the composite, a grade or classification is then assigned. Generally, separate standards, or subsections of, address classification of the various nonstructural wood composites. Tables 2–4 list some of the characteristics by which the American National Standards Institute (ANSI), Joint Standards Australia/New Zealand Committee (AS/NZS) and the Comité Européen de Normalisation (EN), assess particleboard, MDF, and hardboard performance. The values listed represent the required performance of the highest-grade panel (standard application) endorsed by each particular organization. The assessment of decorative plywood is generally by both qualitative and quantitative methods, the diversity of which is beyond the scope of this review.

As both composite production and consumption becomes ever more international, the importance of unifying classifying standards has become paramount. At the time of writing, international unification of nonstructural wood-composite standards is proposed under the auspices of the International Organization for Standardization, while unification on a lesser level, namely the European Community, has already been established.

See also: Wood: Structural Panels*; Wood: Materials for Veneers and Panels$^{\#}$; Wood: Nonstructural Panel Processes*

Bibliography

Composite Panel Association 1999–2000 *CPA Buyers and Specifiers Guide.*

Dinwoodie J M 1981 *Timber: Its Nature and Behavior*. Van Nostrand Reinhold, New York

Heslop G 1997 Compak: Ten Years of Experience with Commercial Straw Particleboard Production. In: Wolcott M P (ed.) *Proc. 31st Int. Particleboard/Composite Materials Symp.* Washington State University, Pullman, WA, pp. 109–14

Maloney T M 1997 *Modern Particleboard & Dry-Process Fiberboard Manufacturing*. Miller Freeman, San Francisco

Sellers T, Haupt R A 2000 New Developments in Wood Adhesives and Gluing Processes in North America. In: *Working Unit 5.04.07 Adhesives and Wood Gluing*. International Union of Forestry Research Organizations, Division 5 Tampere, Finland (http://iufro.boku.ac.at/iufronet/d5/wu50407/pub/finland.htm)

T. Adcock and M. P. Wolcott
Washington State University, Pullman, Washington, USA

Wood: Strength and Stiffness

Wood is one of the oldest and best-known structural materials, and one of the few renewable natural resources. Wood is a desirable material for construction because it requires less energy to produce a usable end product than do other materials. Wood is also extremely versatile. Because wood has a wide range of physical and mechanical properties, species can be selected on the basis of how well they fit the requirements for a particular product. The ability of wood to resist loads depends on a number of factors, including the type, direction, and duration of loading; ambient conditions of moisture content and temperature; and the presence or absence of defects such as knots and splits. This article is concerned with the properties of clear wood and the factors that influence these properties.

1. Orthotropic Nature of Wood Properties

The cellular structure of wood and the physical organization of the cellulose chain within the cell wall make the physical and mechanical properties of wood dependent upon the direction of loading (see *Wood: Ultrastructure**). Wood may be described as an orthotropic material; that is, it has unique and independent mechanical properties in the directions of three mutually perpendicular axes (Fig. 1). The longitudinal axis L is parallel to the cylindrical trunk of the tree and therefore also to the long axis of the wood fibers (parallel to the grain). The tangential axis T is perpendicular to the grain but tangent to the annual growth rings, and the radial axis R is normal to the growth rings. Collectively, the tangential and radial directions are referred to as being perpendicular to the grain. The properties of wood parallel to the grain are higher than those perpendicular to the grain, since the grain direction is also the direction of the primary bonds of the major chemical constituents of the wood cell wall.

1.1 Elastic Properties

The elastic properties of wood are those produced at low stress levels and are completely recoverable after the loads are removed. Twelve constants, nine of which are independent, are needed to describe the elastic behavior of wood: three moduli of elasticity (E), three moduli of rigidy (G), and six Poisson ratios (μ). The three moduli of elasticity are denoted by E_L, E_T, and E_R for the three orthotropic axes. These moduli are used to characterize the unit deformation (or strain) in the orthotropic directions, and are usually determined from compression tests. However, data for E_T and E_R are not extensive. Average values for a few species are presented in Table 1 as ratios with E_L. The modulus of elasticity determined by bending (E_B) rather than by axial load may be the only value available for a given species. Average values of E_B determined by bending tests are given in Table 2 for selected species. As tabulated, these E_B values include an effect of shear deflection. The E_B values determined from bending tests are usually increased by 10% when estimating E_L for axial tests.

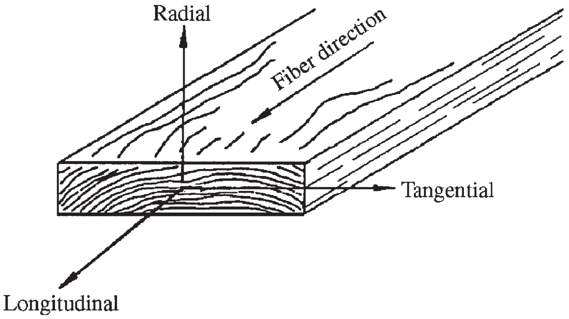


Figure 1
Three principal axes of wood with respect to grain direction and growth rings.

Values for the moduli of rigidity, also called shear moduli, are also given in Table 1 as ratios with E_L. The three shear moduli are denoted G_{LR}, G_{LT}, and G_{RT}, where the subscripts refer to the plane over which the shear strain is measured. When a member is loaded axially, the deformation perpendicular to

Table 1
Major elastic constants for five wood species at 12% moisture content.[a]

Property	Loblolly pine	Sitka spruce	Red oak	Yellow poplar	Balsa
E_T/E_L	0.078	0.043	0.082	0.043	0.015
E_R/E_L	0.113	0.078	0.154	0.092	0.046
G_{LR}/E_L	0.082	0.064	0.089	0.075	0.054
G_{LT}/E_L	0.081	0.061	0.081	0.069	0.037
G_{RT}/E_L	0.013	0.003	—	0.011	0.005
μ_{LR}	0.328	0.372	0.350	0.318	0.229
μ_{LT}	0.292	0.467	0.448	0.392	0.488
μ_{RT}	0.382	0.435	0.560	0.703	0.665
μ_{TR}	0.362	0.245	0.292	0.392	0.231

[a] Values for μ_{RL}, and μ_{TL} are small, seldom used, and often not available. Values of E_L may be estimated by multiplying the modulus of elasticity in static bending given in Table 2 by 1.10.

Table 2
Mechanical properties for five wood species at 12% moisture content.[a]

	Bending		Compression		Shear	Tension		
Species	Modulus of rupture (MPa)	Modulus of elasticity (GPa)	Parallel to grain (MPa)	Perpendicular to grain (MPa)	Parallel to grain (MPa)	Parallel to grain (MPa)	Perpendicular to grain (MPa)	Side hardness (kN)[a]
Loblolly pine	88.0	12.3	49.2	5.4	9.6	88.0	3.2	3.1
Sitka spruce	70.0	10.8	38.7	4.0	7.9	75.8	2.6	2.4
Red oak	99.0	12.5	46.6	7.0	12.3	101.4	5.5	5.7
Yellow poplar	70.0	10.9	38.2	3.4	8.2	154.4	3.7	2.4
Balsa	21.6	3.4	14.9	—	2.1	—	—	—

[a] Force at 5.6mm indention.

the direction of loading is proportional to the deformation in the direction of loading. The constants that characterize this proportionality are called Poisson ratios, and are denoted μ_{LR}, μ_{LT}, and μ_{RT}. Here, the first subscript refers to the direction of applied load and the second subscript to the direction of lateral deformation. The elastic constants, as well as the ratios of elastic constants, vary by species and the moisture content and temperature at which they are measured.

1.2 Strength Properties

When wood is loaded to higher stress levels beyond the elastic range, plastic deformation or failure occurs. Five strength properties that are commonly measured for design purposes include bending, compression parallel and perpendicular to the grain, tension parallel to the grain, and shear parallel to the grain. In addition, measurements are sometimes required for tensile strength perpendicular to the grain and side hardness. Strength data for clear, defect-free specimens for a few species are given in Table 2. Procedures for making these measurements using clear, straight-grained material are given in ASTM standard D143.

For clear defect-free wood, the bending test probably provides the most widely available strength property. The *modulus of rupture* (MOR) reflects the maximum load-carrying capacity in bending and is proportional to the maximum moment borne by the specimen. Modulus of rupture is an accepted criterion of strength, although it is not a true stress because the formula by which it is calculated is valid only within the elastic range. Compression strength parallel to the grain is much lower than tensile strength. During a bending test, initial yielding occurs on the compression side, followed by visible compression failures and enlargement of the compression zone. The neutral surface shifts toward the tensile side of the specimen as the tensile stress continues to increase. The maximum moment in the member is reached when failure in tension occurs. The MOR of clear wood is intermediate between tensile and compression strength parallel to the grain.

Ultimate compression stress parallel to the grain, UCS, when tested as a short column, is initially characterized by folding of the cellulose microfibrils. This folding can begin to occur at stress levels as low as one-half that of the ultimate failure stress. At higher stress levels, a similar folding takes place at the cell wall level, and eventually leads to gross failure of the specimen. In *compression perpendicular to the grain*, C-perp, failure occurs by collapse and flattening of the wood cells. During testing, clearly defined maximum load cannot usually be obtained, and only fiber stress at the proportional limit or stress at a defined deformation limit is reported. Although C-perp values do not show a particular maximum for either the radial or tangential orientation, values are often found to be a minimum at 45° to the annual growth increment for softwood species.

Ultimate tensile stress parallel to the grain, UTS, is difficult to obtain experimentally with clear defect-free wood. The value parallel to the grain is of the order of 45–120 MPa at 12% moisture content, whereas the *tensile stress perpendicular to the grain*, T-perp, may only be 2–6% of the parallel-to-grain value. Thus, it is difficult to get wood to fail in tension parallel to the grain without having excessive failure in tension perpendicular to the grain. For this reason, only a limited amount of data is available on the tensile strength of clear wood parallel to the grain. Values for T-perp are determined as an average of the values in the radial and tangential directions.

Shear strength parallel to the grain ranges from 3 to 15 MPa at 12% moisture content. Because wood is highly orthotropic, it is very difficult to get it to fail in shear perpendicular to the grain. Attempts to obtain shear failure perpendicular to the grain usually result in failure in another failure mode, such as compression perpendicular to the grain. A very limited amount of data suggests that shear strength perpendicular to the grain may be 2.5–3 times that of shear parallel to the grain.

2. Factors Affecting Wood Properties

2.1 Natural Characteristics Related to Wood Structure

Much of the variation in wood properties within and between trees can be attributed to density. The cell wall substance is actually heavier than water; with a specific gravity of about $1.5\,\mathrm{kg\,m^{-3}}$ regardless of species. The dry wood of most species nevertheless floats in water, and thus it is evident that part of the volume of a piece of wood is occupied by cell cavities and pores. Variations in the size of these openings and in the thickness of the cell walls cause some species to have more wood substance per unit volume than do other species. In the absence of knots and other defects, specific gravity is therefore an excellent indicator of the amount of wood substance present and is a good predictor of mechanical properties. Approximate relationships between various mechanical properties and specific gravity for clear straight-grained wood are given in Table 3.

A *knot* is that portion of a branch that has become incorporated in the bole of a tree. The influence of a knot on mechanical properties is primarily due to the interruption of continuity and change in direction of wood fibers around the knot. The influence of a knot on the performance of lumber depends upon the size, location, and shape of the knot as well as attendant grain deviation around the knot and the type of stress to which the wooden member is subjected. Knots

Table 3
Functions relating mechanical properties to specific gravity of clear straight-grained wood.[a]

	G (ρ–strength relationship)	
Property[b]	Softwoods	Hardwoods
Static bending		
Modulus of rupture (MPa)	$170.7G^{1.01}$	$171.3G^{0.13}$
Modulus of elasticity (GPa)	$20.5G^{0.84}$	$16.5G^{0.7}$
Compression parallel (MPa)	$93.7G^{0.97}$	$76.0G^{0.89}$
Compression perpendicular (MPa)	$16.5G^{1.57}$	$21.6G^{2.09}$
Shear parallel (MPa)	$16.6G^{0.85}$	$21.9G^{1.13}$
Tension perpendicular (MPa)	$6.0G^{1.11}$	$10.1G^{1.3}$
Side hardness (kN)	$85.9G^{1.5}$	$15.3G^{2.09}$

[a] Wood at 12% moisture content. [b] Compression parallel to grain is maximum crushing strength; compression perpendicular to grain is fiber stress at proportional limit. Specific gravity based on oven dry mass and volume at 12% moisture content (USDA 1999).

have a much greater effect on strength in axial tension than in axial compression. The effect on bending is somewhat less than that in axial tension.

In some wood products the direction of critical stresses may not coincide with the orthotropic axes of the material. This may occur by choice in design, or it may be a result of the way the wood was removed from the log. This *cross-grain* can have a major effect on mechanical properties. Elastic properties in directions other than those along the orthotropic axes can be obtained from elastic theory. Strength properties in directions other than parallel (P) or perpendicular (Q) to the grain can be approximated using a Hankinson-type formula:

$$N = \frac{PQ}{P \sin^n \theta + Q \cos^n \theta}$$

where θ is the angle to the grain. For bending and tensile strength, N ranges from 1.5 to 2, and for compressive strength from 1.5 to 2.0. For the modulus of elasticity an N value of 2 is often recommended.

2.2 *Effects of Manufacturing and Use Environments*

Many mechanical properties are affected by changes in *moisture content* below the fiber saturation point. Generally, most mechanical properties increase as wood is dried. The strength P at moisture content M in the range from 8 to 25% can be estimated by the following equation:

$$P = P_{12}\left(\frac{P_{12}}{P_g}\right)^{(12-M)/(M_p-12)}.$$

For the purposes of the formula, moisture content of green wood (M_p) is often assumed to be 25% for most species. ASTM D2555 gives "dry/green" ratios of mechanical properties at 12% moisture content compared with green wood. Care must be taken if this formula is used at moisture contents very far below 12%. At low moisture content levels, the rate of change of properties becomes less than that predicted by the equation, and for some properties there may even be a maximum value with loss in property during further drying. Above the fiber saturation point, most mechanical properties are not affected by change in moisture content.

Temperature can have both immediate (reversible) and permanent (irreversible) effects on wood properties (see *Wood Products: Thermal Degradation and Fire*[#]). In general, one immediate effect is that mechanical properties tend to decrease as the temperature is increased (Fig. 2). There is an interaction with moisture content because dry wood is less sensitive to temperature change than is green wood. However, increases in temperature are usually accompanied by a reduction in moisture content. Permanent loss in mechanical properties can occur if wood is subjected to high temperatures over long periods. The magnitude of this effect depends upon temperature, duration of exposure, wood moisture content, and wood property. Permanent effects resulting from temperature are discussed in a report by the US Department of Agriculture Forest Service (1999).

In warm, moist conditions wood may be subject to *decay*. Decay fungi remove wood substance, thus lowering the gross specific gravity of the material. Strength losses can occur rapidly, even during the early stages of decay before noticeable loss in mass has occurred (see *Wood Products: Decay During Use**).

Mechanical properties, as given in Table 2, are usually referred to as static strength values. The tests of such properties are commonly conducted at a rate of loading such that failure occurs in 5–20 min. Higher values are obtained when wood is loaded

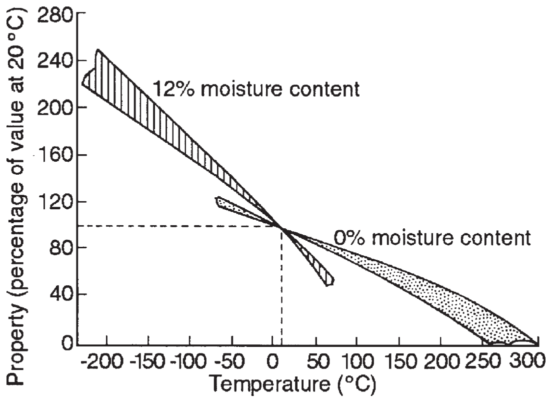


Figure 2
Immediate effect of temperature on bending strength, tensile strength perpendicular to grain, and compression strength parallel to grain. Variability is indicated by the width of bands.

faster, and lower values when it is loaded slower. Load capacity decreases from the static values by about 7–8% for each decade increase on a logarithmic time-scale. Thus, a wooden member subjected to a continuous bending load for 10 years may carry only 60% or less of the load required to produce failure in a static test. *Creep* is additional time-dependent deflection over that resulting from elastic deformation which occurs when wood is subjected to a constant load over a period of time. Changes in climatic conditions during a duration of load or creep test may produce a lower load capacity and more creep than that observed under constant conditions of temperature and moisture content. These effects can be quite substantial for small wood specimens. Fortunately, such changes are moderate for most wood structural members in typical service environments.

2.3 Allowable Engineering Properties of Lumber

When sawn, lumber yields wood of varying quality. Grading rules quantify the quality of lumber required for specific uses. Testing procedures for lumber and procedures for assigning allowable properties to the structural grades of lumber are given in ASTM standards such as D4761, D245, and D1990. In the USA, the National Design Specifications summarize allowable properties for structural grades of lumber and provide adjustment procedures for environmental and end-use conditions. The latest versions of national design specifications should be consulted when designing wood structures.

See also: Wood: Density[#]; Wood: Creep and Creep Rupture*; Wooden Structures[#]

Bibliography

American Forest and Paper Society 1997 *National Design Specifications for Wood Construction.* American Forest and Paper Society, Washington, DC

ASTM 1999 *Annual Book of Standards*, Vol. 04.10. *Wood.* American Society for Testing and Materials, West Conshohocken, PA

Bodig J, Jayne B A 1982 *Mechanics of Wood and Wood Composites.* Van Nostrand Reinhold, New York

Green D W, Hernandez R 1998 Standards for wood products and their use in the United States. *Wood Design Focus* **9** (3), 3–12

USDA Forest Service 1999 Wood handbook. Wood as an engineering material. General Technical Report FPL–113. USDA Forest Service, Forest Products Laboratory, Madison, WI

D. W. Green
USDA Forest Service, Madison, Wisconsin, USA

Wood: Structural Panel Processes

This article discusses the processes used to produce oriented strandboard (OSB) and structural plywood (for information on particleboard, fiberboard, and hardboard see *Wood: Nonstructural Panel Processes**).

The manufacturing process is designed to create structural panels by converting logs into veneer or flakes, combining these wood particles with adhesive, wax, and other additives, arranging them in a loose mat, and finally consolidating this mat under heat and pressure to form a large panel. Panel sizes can range from 1220 mm to 3660 mm wide and from 2440 mm to 8534 mm long. Regardless of whether the product is OSB or plywood the unit operations are basically the same although the equipment is quite different. These operations can be grouped as follows:

(i) the "green end" consisting of wood receiving, handling, sorting, preparing, and flaking or peeling;
(ii) particle drying (removing moisture);
(iii) applying adhesive to veneers or flakes;
(iv) forming a loose mat of flakes or veneers; and
(v) pressing the mat using heat and pressure to create a structural panel.

1. Raw Material: Wood

Wood makes up more than 90% of the panel. Therefore, the species available and the quality of the resource are very important to the successful operation of the manufacturing facility. Structural plywood and OSB can be made from many species although not all species will produce panels of equal properties.

The success of the manufacturing process is influenced strongly by raw material quality. Each mill sets requirements regarding the size of logs, the allowable species, the form such as crook, sweep, knots, and the amount of degrade such as decay that will be allowed. These requirements will have an impact on the efficiency of an operation. For instance, small diameter crooked logs or logs with considerable decay will increase waste and reduce the yield of on-grade veneer or flakes. Under these circumstances, manufacturing a given quantity of finished product will require handling more logs and disposing of more waste, thus increasing costs compared to using only high-quality, sound logs. In areas where there is little competition for wood, a mill will be able to set and maintain strict limitations on the quality of wood that will be accepted. However, in many areas where there is strong competition for wood, a mill may have little or no choice in the quality of wood that it receives unless considerably more is paid for it.

1.1 Woodyard

The purpose of the woodyard is to receive, inspect, and place into storage logs that meet the quality requirements of the mill. Wood is obtained either from company-owned lands or purchased from independent contractors who log public and private forests. Depending on the area traditions or the design of the mill, logs are purchased as pulpwood (about 2540 mm long) or tree length. Wood is received at the truck scale, and is purchased either by weight or volume. If it is weighed many mills will use seasonal moisture content factors which account for the fluctuations in moisture content of the wood throughout the year. The truck scale is usually the first point of inspection and the first opportunity to reject wood that does not meet quality standards. Once accepted the logs are stored. Some mills (especially plywood mills) will store logs in ponds or spray water onto the logs during storage to reduce degrade such as blue stain and to prevent the logs from drying. This is rarely done in the OSB industry. Most woodyards are not paved, although paving reduces the amount of sand, rocks, grit, and other debris carried into the process. Paving reduces the frequency of cleaning the conditioning ponds, reduces the maintenance on conveyors, chains, etc., and increases the knife life during the peeling and flaking operations. The rotation out of the woodyard is very important since wood recovery is highest when processing freshly harvested wood. The "first in, first out" scheme in inventory management is desirable.

(*a*) *Species mix.* Logs are either segregated by species or species groups, or well mixed as they are removed from the woodyard. Plywood plants usually process species separately or by groupings based on the strength and stiffness of the wood. The resulting veneer is kept separated throughout the process and the veneer from stronger and stiffer species is used in the surface layers of the panel. OSB plants may segregate and process species for use in the surface or core layers, but this is not common. OSB plants using more than one species must achieve good mixing to ensure consistent properties. This is accomplished as operators remove logs from inventory and put them into the conditioning ponds. If species are not well mixed then the only alternative is to manufacture a product at sufficiently high density and resin content to ensure that the species with the worst properties will perform adequately in the final panel.

1.2 Wood Preparation

The logs must be prepared in such a way that the resulting wood particles—veneers or flakes—will have the required dimensions, the smoothest possible surface, the least amount of cell wall damage, and the least amount of fines (waste particles). Minimizing wood damage is of primary concern in all of the unit operations at the green end of the mill and especially during the conversion of the log into the basic wood element used for panel manufacturing. Turner (1954) clearly demonstrated the effect of particle type and cell wall damage on the final properties of panels. He concluded that the particle shape and cell wall damage are of primary importance and have a greater influence on strength than resin content. Geimer *et al.* (1985) investigated the effect of wood damage in flakes on final properties of flakeboard (OSB) and concluded that flakes damaged during the flaking operation or during hot pressing could reduce significantly tensile strength and internal bond. Therefore, producing high-quality, undamaged veneer or flakes is the most important consideration in the structural panel process. Many people claim that the wood preparation area is where the structural panel is made, and, indeed, errors made in this area can have an enormous impact on profitability. Mistakes in manufacturing that occur in the preparation phase can only be corrected by using additional resin or increasing the density of the panel, which can increase costs significantly.

The most important considerations in converting logs to veneer or flakes include:

(i) Log conditioning or heating the wood to a proper temperature for the species. Wood temperature has a large effect on the quality of the veneer and flake. Surface smoothness, geometry, cell wall damage, and fines generation are all affected by the temperature of the log. If the wood is frozen it is impossible to produce an acceptable veneer or flake due to shattering, wood damage, and fines generation.

(ii) Proper orientation of the log to the knife during peeling or flaking. Whether peeling veneer or cutting flakes or strands it is critical that the log be parallel to the knife and that the grain angle in the resulting wood element is aligned in the long direction of the resulting wood element. Grain deviation through the thickness or in the plane of the veneer or flake should be minimized.

(iii) Proper knife cutting angle. The angle between the log and the knife is critical. Low- to medium-density hardwoods such as aspen will produce excellent flakes at 32–35° cutting angles while softwood species such as southern pine produce better quality flakes using 28–30° cutting angles. Veneers are produced at cutting angles of 18–23°.

(iv) Proper pressure of log against knife. There must be enough pressure to minimize movement between the log and the knife, but not so much that excessive compression of the wood occurs. This will result in excessive thickness variation and cell wall damage (wood failure).

(v) Proper speed. The speed between the log and the knife can influence greatly the flake quality.

Wood preparation begins with log conditioning, especially in cold climates where freezing is common. OSB mills typically have log conditioning ponds or vats to thaw wood. Plywood mills often use steaming chambers or hot water vats to elevate the temperature of the logs. The OSB industry tends to remove bark after the logs have been conditioned whereas the plywood industry usually removes the bark before conditioning the logs. Logs are loaded into one end of the pond or chamber and are discharged from the other after reaching the desired temperature. Each species has an optimum temperature necessary to produce veneer or flakes with acceptable surface quality and strand dimensions. An excellent discussion of proper wood conditioning is given by Lutz (1974). Acceptable veneer or flakes cannot be made from frozen wood. Frozen wood shatters during peeling or flaking. If any acceptable wood elements are produced, the surface quality is extremely poor—rough and uneven. The percentage of on-grade flakes in OSB is greatly reduced when processing frozen wood and the fines content is very high. Simply thawing wood will not ensure that veneer or flakes will be acceptable and Lutz presents a figure showing the optimum temperature for peeling veneer from various species. For instance, quaking aspen (*Populus tremuloides*) should be conditioned to about 21 °C, yellow poplar (*Liriodendron tulipifera*) to 38 °C, and Douglas fir (*Pseudotsuga menziesii*) to about 55 °C.

Logs should be heated uniformly from the outer surface to the core. However, often this is not practical due to the long time required. Therefore, many mills will peel or flake logs when the core temperature is just above freezing (4–5 °C) in the winter to ambient summer temperatures (about 21 °C). Many plywood and OSB mills located in areas where winter temperatures are moderate, and freezing is not common, do not have steaming chambers or log conditioning vats. The reasoning used is that acceptable quality structural panels can be produced even if the wood is cold but not frozen. However, these mills will tend to use more resin or manufacture a higher-density panel to offset the poorer quality of the resulting particle.

Bark is structurally weak compared to wood and is removed from the logs before peeling or flaking. Sand and grit is often embedded in bark, which reduces knife life during peeling or flaking. Debarking is accomplished by using ring debarkers (most common) or large, rotating drum debarkers, and may be done before or after log conditioning. Bark is used as fuel at the mill or sold for other uses. After the logs are properly conditioned and debarked they are peeled or flaked. Although the plywood industry usually maintains the length of veneer (grain direction) at about 2440 mm, the OSB industry generally produces flakes which are 75–150 mm long. Several studies (e.g., Brumbaugh 1960, Geimer and Price 1978) have shown that this is a sufficient length to achieve desired properties when aligned properly.

The structural plywood industry usually produces veneer using rotary lathes. The peeler logs are fixed on each end by chucks and the log is turned against a knife. Lutz (1974) states that knife cutting angle and pressure bar positioning are extremely important in creating high-quality veneer. The only movement should be the log turning against the knife and all other movement between the peeler log and the knife must be minimized. All veneer has tight and loose sides. The loose side refers to the small fractures—lathe checks—which occur on the knife side of the veneer as it is separated from the log. Proper knife grinding, proper knife cutting angle, and proper pressure applied by the pressure bar will ensure high-quality veneer with the least fractures on the loose side and with smooth surfaces. Veneer is fed onto long conveyors, which act as storage to handle the surge of veneer as it exits the lathe. A clipper is located at the end of the conveyor. This is a knife with rapid vertical movement. As the veneer moves beneath the clipper it is activated to cut the long ribbon of veneer into proper widths and to remove defects such as knots and splits.

The ring flaker is the most common machine used by the OSB industry to produce flakes. These machines have a series of knives (30 or more) set into a ring, and the ring and motor are mounted onto a moving carriage. Logs are fed into the center of the ring, held firmly in place and the rotating ring is moved into and through the logs. The knives extend slightly beyond the ring to allow this movement. Flakes pass through the ring and are collected on a conveyor beneath the machine and transported to a green flake storage bin. Proper knife cutting angle and accurate placement of knives into the flaker ring are important in achieving a high yield of good-quality flakes. The carriage hydraulic system must be capable of moving the knife ring into the logs smoothly and at a consistent speed. Any inconsistencies and any movement of the logs will lower the flake quality, increase thickness variation, and increase the percentage of fines. Some manufacturers will remove the fines before the flakes are conveyed to the green bins.

The removal of defects from green veneer by the plywood manufacturer at the clipper and the removal of fines from the green flakes in OSB by screening is an important step in the process. By removing unusable material before drying, dryer capacity and energy is used to remove moisture only from on-grade material, which improves the efficiency of the process. The rejected material can be used as fuel.

2. *Drying Wood Veneer and Flakes*

The purpose of drying is to reduce the moisture content of veneer and flakes. The moisture content of the wood particles entering the dryer can be more than 100% (dry weight basis) and must be reduced to between 2% and 12% moisture content. Drying must be done in a manner that reduces defects such as surface deactivation (resulting in poor adhesive bonds) and buckling. Drying veneer is done using mechanical dryers, which not only transport the veneer through a hot, dry environment, but also apply restraining forces to reduce the degree to which the veneer will buckle. These conveyor dryers have several levels. Each level has closely spaced opposing rollers, which apply the restraining force as the veneer moves between them. Controlling veneer thickness at the lathe (absolute and variability) is critical to successful drying. The areas of greater thickness within a sheet of veneer or between sheets will dry more slowly resulting in higher moisture content. The temperature of the air, velocity of the air impinging on the surface of the veneer, and how much veneer is in the dryer will affect the final moisture content and the variability of moisture content between sheets of veneer. Also, mixing sapwood and heartwood veneer can result in moisture content variability.

The OSB industry usually uses single or triple pass rotary drum dryers to reduce the moisture content of flakes. These dryers are large rotating drums with internal structures enabling flakes to pass through the dryer just once or three times. It appears that selection of single or triple pass dryers is based on personal preference. There is little consensus as to whether one has an advantage over the other either in level of emissions, flake breakage, inlet air temperatures, or final flake moisture content. Conveyor dryers, while not common, are used in some OSB plants. These dryers place flakes on a screen conveyor and transport them through a long multiple-decked dryer. Hot air is forced through the screen and the mat of flakes removing moisture. The manufacturers of these dryers claim less flake breakage, and lower temperature drying resulting in less dryer emissions. Pollution from any dryer can be significant. Manufacturers approach this problem either by lowering the inlet air temperature or by installing end-of-pipe control systems. There is a significant benefit to the former approach in that emissions are reduced by lowering air temperatures, thereby reducing the costs of purchasing, installing, and maintaining pollution abatement devices. In addition, the surface characteristics of the veneer or flake may be improved leading to better adhesive bonding and lower resin costs.

2.1 Dry Veneer and Flake Storage

In plywood manufacturing, veneer exits the dryer and is checked for moisture content. Wet pieces or pieces with significant wet spots may need to be dried further. Veneer is sorted by grade and width. Defects such as knots may be removed and a patch inserted if the grade of the finished panel requires smooth, defect-free surfaces on one or both sides of the panel. In addition, for certain grades or for larger panels (the majority of structural plywood is manufactured in standard 4 ft by 8 ft panels in the USA and in 1200 mm by 2400 mm panels elsewhere) veneer can be edge glued to create a continuous sheet. In the OSB process, dry strands are usually screened to remove fines and then conveyed to a dry flake storage bin. These bins are designed to spread the incoming material so that flakes with small variations in moisture content will not remain together and form a high-moisture pocket. The outlet end of the bin is designed similarly so that strands that passed through the dryer together are not fed out of the bin together.

3. *Applying Adhesive*

The purpose of this operation is to apply the proper amount of adhesive to veneers or flakes. The distribution of adhesive on the surface of the veneer or flakes is a critical factor in the performance of the structural panel. Those surfaces that are to receive adhesive should be uniformly covered with the proper amount of adhesive.

Structural panels are required to be glued with adhesives that are water resistant. The primary adhesive in both the plywood and OSB industries is phenol formaldehyde (phenolic), which is synthesized from phenol and formaldehyde. In addition, the OSB industry uses methylene biphenyl isocyanate (also known as isocyanate or MDI). These adhesives cure rapidly when subjected to elevated temperatures during the pressing operation. There are a variety of means available to alter the rate of cure, so that adhesives used in the core of OSB, for instance, will cure more rapidly and at lower temperatures than those used in the surface layers.

Plywood manufacturers use extenders to improve the flow of adhesive on the surface of the veneer and to increase the open assembly time (the time between the veneer assembly and when it is pressed). Adhesives can be modified to allow for differences related to species and seasonal changes.

Roll spreaders, spray lines, curtain coaters, and foam glue extruders apply phenolic adhesives to sheets of veneer in plywood manufacturing. The choice of equipment is dependent on the design and capacity of the process. The OSB industry primarily uses large rotating drums within which are either spinning disks or spray nozzles. These drums are angled and flakes are conveyed into one end. Adhesive is pumped through spray nozzles or onto the surface of a spinning disk. Both devices create droplets, which are discharged towards the mass of flakes. In both plywood and OSB manufacturing the accurate control of wood and adhesive is necessary to meet the required product specifications and to minimize variability within and between panels.

4. *Mat Formation*

In plywood manufacturing, the veneers with adhesive are quickly assembled into a loose mat. The proper grade of veneer must be placed in the appropriate position within this assembly. Veneers of lower strength or with more or larger defects can be placed in the core while the highest strength veneers with the fewest defects are placed on the surface. Plywood alternates the grain direction of adjacent plies. Surface veneers have grain running in the long direction of the panel. The plies just below the surface veneers have the grain running at right angles to the long direction. This alternating grain direction continues throughout the thickness. Linear expansion, strength, and stiffness are improved in the cross-panel direction while cupping, twisting, and warping are reduced by alternating the grain direction of the veneers throughout the thickness of plywood.

In OSB manufacturing, flakes are discharged from the blender and fed into four or more forming bins, which are positioned in a linear sequence above a long, wide conveyor (forming line). Flakes are discharged carefully from each former and must pass through orienters before falling onto the moving forming line belt. The first and last formers create the two surfaces of the mat and align the flakes in the long direction of the panel. These machine direction orienters are usually rotating shafts onto which are mounted closely spaced disks. Flakes are cast onto the rotating disks and must orient themselves lengthwise in order to fall through and onto the forming line. The formers between the surface formers create the cross-oriented core layers. The flakes are cast onto a finned aligning device, which causes the strands to turn in the cross-panel direction before falling onto the forming line.

5. *Pressing*

The purpose of this operation is to consolidate the loose mat of veneer or flakes as quickly as possible, elevate the temperature of the mat, and to cure the adhesive as quickly as possible, so transforming the mat into a structural panel. The plywood and OSB processes usually use multiple-opening presses with fixed platen dimensions. These may have up to 50 openings although most have between 12 and 24 openings. The size of the platens can be as small as about 1220 mm by 2440 mm to as large as 3660 mm by 7320 mm, and at least one is longer. One continuous press is used to manufacture OSB. This concept reduces waste (eliminates end trimming) and provides much more flexibility in finished panel size, which is limited only by handling equipment and shipping restrictions.

Both plywood and OSB technologies may use a prepress to consolidate the mat prior to loading into the hot press. This technique presses the surfaces of veneers or flakes into intimate contact with each other using low pressure. The purpose is to prevent drying of the adhesives and improve handling of the mat. This process also reduces the thickness of the mat making the hot-pressing operation more efficient. Plywood manufacturing would typically use a platen-type prepress (cold platens) and the OSB process would use continuous belt prepresses. Mats are then transported to a loading cage (in the case of multiple-opening presses) and then an entire press load is put into the hot press with all openings being filled at the same time. Since the platen temperature of these presses is maintained between 150 °C to as high as 210 °C, it is critical that the press closes rapidly and consolidates the mat to the final panel thickness before the adhesive begins to cure. If the press delays in closing then precuring of the adhesive will result in poorly bonded panels with poor durability.

6. *Panel Processing*

As the finished panels exit the press they are trimmed to rough dimensions. Panels are then usually sent directly to a hot stack where the adhesive finishes curing and properties improve. After a suitable duration the panels are cut to exact finished sizes and prepared for shipping.

Acknowledgment

I would like to thank Dr Michael R. O'Halloran for his assistance in supplying information about the plywood manufacturing process.

See also: Wood: Materials for Veneers and Panels$^{\#}$; Wood: Nonstructural Panel Processes*; Wood: Structural Panels*; Wooden Structures$^{\#}$

Bibliography

Brumbaugh J 1960 Effect of flake dimensions on properties of particleboard. *For. Prod. J.* **10** (5), 243–6

Geimer R L, Mahoney R J, Loehnertz S P, Meyer R W 1985 Influence of processing-induced damage on strength of flakes and flakeboards. USDA Forest Service Research Paper FPL 463

Geimer R L, Price E W 1978 Construction variables considered in fabrication of a structural flakeboard. USDA Forest Service General Technical Report WO-5

Lutz J F 1974 Techniques for peeling, slicing, and drying veneer. USDA Forest Service Research Paper FPL 228

Turner H D 1954 Effect of particle size and shape on strength and dimensional stability of resin-bonded wood-particle panels. *J. For. Prod. Res. Soc.* (October), 210–23

A. W. Boehner
Trus Joist – A Weyerhaeuser Business, Boise, Idaho, USA

Wood: Structural Panels

Structural-use panels is a term used to identify a class of utilitarian wood-based panel products used for construction and industrial applications. The primary market for structural wood panel products is the light-frame construction industry. Panel products are shipped to suppliers and contractors for on-site assembly of residential and commercial structures. These wood panels are manufactured and distributed in standard sizes and grades.

Structural panel products include plywood, oriented strandboard (OSB), and waferboard. Plywood is one of the oldest ''engineered wood'' products, and has a long history of application over a broad range of uses. OSB and waferboard began emerging in the late 1970s, and now represent over 50% of the structural-use panels in production in North America.

1. Plywood

Plywood comprises a number of thin layers of wood called veneers, which are bonded together by an adhesive. Each layer is placed so that its grain direction is at right angles to that of the adjacent layer. This cross-lamination gives plywood its characteristics and makes it a versatile building material. Plywood has been used for centuries; it has been found in Egyptian tombs and was in use during the height of the ancient Greek and Roman civilizations.

The modern plywood industry is divided into the so-called hardwood and softwood industries. While the names are truly misnomers, they are in common use, the former referring to industry manufacturing decorative paneling and the latter referring to plywood for building construction and industrial uses. This article is concerned primarily with the plywood manufactured for construction and for industrial uses in North America.

1.1 Standards

Construction plywood in the USA is manufactured under US Product Standard PS 1 for Construction and Industrial Plywood (US Department of Commerce 1995) or a Performance Standard such as those promulgated by the APA–The Engineered Wood Association, formerly the American Plywood Association. The Product Standard for plywood includes provisions for performance rating as well. The standard for construction and industrial plywood includes over 70 species of wood, which are separated into five groups based on their mechanical properties. This grouping results in fewer grades and a simpler procedure.

Certain panel grades are usually classified as sheathing and include C-D, C-C, and Structural I C-D and C-C. These panels are sold and used in the unsanded or ''rough'' state, and are typically used in framed construction, rough carpentry, and many industrial applications. The two-letter system refers to the face and back veneer grade. A classification system provides for application as roof sheathing or subflooring. A fraction denotes the maximum allowable support spacing for rafters (in inches) and the maximum joist spacing for subfloors (in inches) under standard conditions. Thus, a 32/16 panel would be suitable over roof rafter supports at a minimum spacing of 32 in (800 mm) or as a subfloor panel over joists spaced 16 in (400 mm) on center. Common ''span ratings'' include 24/0, 24/16, 32/16, 40/20, and 48/24.

Many grades of panels must be sanded on the face and/or back to fulfill the requirements of their end use. Sanding most often takes place on plywood panels with A- or B-grade faces and some C ''plugged'' faces are sanded as well. Sanded panels are graded according to face veneer grade, such as A-B, A-C, or B-C. Panels intended for concrete forming are most often sanded as well and are usually B-B panels. Panel durability for plywood is most often designated as Exterior or Exposure 1. Exterior plywood consists of a higher solids-content glueline and a minimum veneer grade of C for any ply of the panel. Such panels can be permanently exposed externally without fear of deterioration of the glueline. Exposure 1 panels differ in the fact that D-grade veneers are allowed. Exposure 1 panels are most often made with an extended phenolic glueline and are suitable for protected exposure and short exterior exposure during construction delays.

1.2 Physical Properties

The plywood described here is assumed to be manufactured in accordance with US Product Standard PS 1, Construction and Industrial Plywood (US Department of Commerce 1983). The physical property data were collected over a period of years (O'Halloran 1975).

(*a*) *Effects of moisture content.* Many of the physical properties of plywood are affected by the amount of moisture present in the wood. Wood is a hygroscopic material, which nearly always contains a certain amount of water. When plywood is exposed to a constant relative humidity (RH) it will eventually reach an equilibrium moisture content (EMC). The EMC of plywood is highly dependent on RH, but is essentially independent of temperature between 0 °C and 85 °C. Examples of values for plywood at 25 °C include 6% EMC at 40% RH, 10% EMC at 70% RH, and 28% EMC at 100% RH.

Plywood exhibits greater dimensional stability than most other wood-based building products. Shrinkage of solid wood along the grain with changes in moisture content is about 2.5–5% of that across the grain. The tendency of individual veneers to shrink or swell crosswise is restricted by the relative longitudinal stability of the adjacent plies, aided also by the much greater stiffness of wood parallel as compared to perpendicular to the grain. The average coefficient of hygroscopic expansion or contraction in length and width for plywood panels with about the same amount of wood in parallel and perpendicular plies is about 0.002 mm mm^{-1} for each 10% change in RH. The total change from the dry state to the fiber saturation point averages about 0.2%. Thickness swelling is independent of panel size and thickness of veneers. The average coefficient of hygroscopic expansion in thickness is about 0.003 mm mm^{-1} for each 1% change in moisture content below the fiber saturation point.

The dimensional stability of panels exposed to liquid water also varies. Tests were conducted with panels exposed to wetting on one side as would be typical on a rain-delayed construction site (Bengelsdorf 1981). Such exposure to continuous wetting on one side of the panel for 14 days resulted in about 0.13% expansion across the face grain direction and 0.07% along the panel. The worst-case situation is reflected by testing from oven-dry to soaking in water under vacuum and pressure conditions. Results for a set of plywoods similar to those tested on one side showed approximately 0.3% expansion across the panel face grain and 0.15% expansion along the panel. These can be considered the theoretical maximum that any panel could experience.

(*b*) *Thermal properties.* Heat has a number of important effects on plywood. Temperature affects the EMC and the rate of absorption and desorption of water. Temperatures below 90 °C have a limited long-term effect on the mechanical properties of wood. Very high temperature, however, will weaken the wood.

The thermal expansion of wood is smaller than swelling due to absorption of water. Because of this, thermal expansion can be neglected in cases where wood is subject to considerable swelling and shrinking. It may be of importance only in assemblies with other materials where moisture content is maintained at a relatively constant level. The effect of temperature on plywood dimensions is related to the percentage of panel thickness in plies having grain perpendicular to the direction of expansion or contradiction. The average coefficient of linear thermal expansion is about 6.1×10^{-6} mm^{-1} °C^{-1} for a plywood panel with 60% of the plies or less running perpendicular to the direction of expansion. The coefficient of thermal expansion in panel thickness is approximately 28.8×10^{-6} mm^{-1} °C^{-1}.

The thermal conductivity, k, of plywood is about 0.11–0.15 W m^{-1} K^{-1} depending on species. This compares to values (in W m^{-1} K^{-1}) of 391 for copper (heat conductor), 60 for window glass, and 0.04 for glass wool (heat insulator).

From an appearance and structural standpoint, unprotected plywood should not be used at temperatures exceeding 100 °C. Exposure to sustained temperatures higher than this will result in charring, weight loss, and permanent strength loss.

(*c*) *Permeability.* The permeability of plywood is different from solid wood in several ways. The veneers from which plywood is made generally contain lathe checks from the manufacturing process. These small cracks provide pathways for fluids to pass by entering through the panel edge. Typical values for untreated or uncoated plywood of ṁ in (9.5 mm) thickness lie in the range 0.014–0.038 g m^{-2} h^{-1} mm Hg^{-1}.

Exterior-type plywood is a relatively efficient gas barrier. Gas transmission ($\times 10^{-5}$ cm^3 s^{-1} cm^{-2} cm^{-1}) for ṁ in (9.5 mm) exterior-type plywood is as follows: oxygen, 2.9; carbon dioxide, 2.6; nitrogen, 2.1.

1.3 Mechanical Properties

The first design document for plywood was published by the APA–The Engineered Wood Association (APA 1986). The woods used to manufacture plywood under US Product Standard PS 1 are classified into five groups based on elastic modulus, and bending and other important strength properties. The 70 species are grouped according to procedures in ASTM D2555 (American Society for Testing and Materials 1987). Design stresses are published only for groups 1–4 since group 5 is a provisional group with little or no actual production. APA (1986) provides for development of plywood sectional properties based on the geometry of the lay-up and the species, and combining these with the design stresses for the appropriate species group. These data have been updated (APA 1999b).

(*a*) *Section properties.* Plywood section properties are computed according to the concept of transformed sections to account for the difference in stiffness parallel and perpendicular to the grain of any given ply. Published data take into account all possible manufacturing options under the appropriate standard, and consequently the resulting published value tends to reflect the minimum configuration. These "effective" section properties computed by the transformed section technique take into account the orthotropic nature of wood, the species group used in the outer and inner plies, and the manufacturing variables provided for each grade. The section properties presented are generally the minimums that can be expected.

Because of the philosophy of using minimums, section properties perpendicular to the face grain direction are usually based on a different configuration from those along the face grain direction. This compounding of minimum sections typically results in conservative designs. Information is available for optimum designs where required.

(*b*) *Design stresses.* Design stresses include values for each of four species groups and one of three grade stress levels. Grade stress levels are based on the fact that bending, tension, and compression design stresses depend on the grade of the veneers. Since veneer grades A and natural C are the strongest, panels composed entirely of these grades are allowed higher design stresses than those of veneer grades B, C-plugged, or D. Although grades B and C-plugged are superior in appearance to C, they rate a lower stress level because the "plugs" and "patches" which improve their appearance reduce their strength somewhat. Panel type (interior or exterior) can be important for bending, tension, and compression stresses, since panel type determines the grade of the inner plies.

Stiffness and bearing strengths do not depend on either glue or veneer grade but on species group alone. Shear stresses, however, do not depend on grade, but vary with the type of glue. In addition to grade stress level, service moisture conditions are also typically presented—for dry conditions, typically moisture contents less than 16%; and for wet conditions at higher moisture contents.

Allowable stresses for plywood typically fall in the same range as for common softwood lumber, and when combined with the appropriate section property, result in an effective section capacity. Some comments on the major mechanical properties with special consideration for the nature of plywood are given below.

Bending modulus of elasticity values include an allowance for an average shear deflection of about 10%. Values for plywood bending stress assume flat panel bending as opposed to bending on edge, which may be considered in a different manner. For tension or compression parallel or perpendicular to the face grain, section properties are usually adjusted so that allowable stress for the species group may be applied to the given cross-sectional area. Adjustments must be made in tension or compression when the stress is applied at an angle to the face grain.

Shear through-thickness stresses are based on common structural applications such as plywood mechanically fastened to framing. Additional options include plywood panels used as the webs of I-beams. Another unique shear property is that termed rolling shear. Since all of the plies in plywood are at right angles to their neighbors, certain types of loads subject them to stresses which tend to make them roll, as a rolling shear stress is induced. For instance, a three-layer panel with framing glued on both faces could cause a cross-ply to roll across the lathe checks. This property must be taken into account with such applications as stressed-skin panels.

2. *Structural-use Panels*

With the advent of waferboard and OSB, the term structural-use panels became commonly applied to panel products used in construction. Such common uses include wall, roof, and floor sheathing. Standards, code recognition, and end-use recommendations drive the acceptance of these products in the marketplace.

2.1 *Performance Standards*

Historical standards against which wood-based panel products have been manufactured have been product standards such as PS 1 for construction and industrial plywood and ANSI A208.1 for mat-formed wood particleboard. National consensus or product standards are most frequently promulgated through organizations such as the American National Standards Institute (ANSI) or, in the case of US Product Standards, the US Department of Commerce. Product or commodity standards are essentially manufacturing prescriptions for minimum-quality products. Requirements are not centered on the specific manufacturing process, and typically do not address the question of product application or how well the product is suited for any particular application.

A performance standard reverses the standard process by defining the end-use of the product and does not prescribe how the product will be manufactured. The objective of a performance standard is to ensure that the product will satisfy the requirements of the application for which it is intended. To do this the performance standard must define performance criteria and test methods. Performance standards were pioneered by the APA–The Engineered Wood Association in 1980, and have since been adopted by other trademarking organizations in the USA, and by the Canadian Standards Association in a standard adopted in 1988. The standards developed by the APA include one for single-layer floors and for sheathing. The intended end-use of these products is for light-frame wood construction as practiced in the USA. As the name implies, panels are intended for combination subfloor underlay and sheathing intended for roofs, walls, and subfloors. The products are rated for their particular end-use. The standards cover panels for supports spaced 400–1220 mm apart for floors and roofs. Provisions are also made for wall bracing when the studs (vertical framing elements) are spaced 400 mm or 600 mm apart. A recent development includes a similar standard for siding where the panels (or panels cut into strips for lap) are to be the exterior building surface.

Panel performance is evaluated through testing under a qualification procedure for three basic areas of investigation: structural adequacy, dimensional stability, and bond performance.

Testing for structural adequacy includes:

(i) verification of the ability to sustain concentrated loads, both static and by impact;

(ii) uniform-load testing to demonstrate the ability of the product to sustain wind or snow load;

(iii) testing of fastener-holding ability to ensure that the panel has the ability to hold covering materials in place and to develop lateral loads in mechanically fastened connections to lumber or other framing; and

(iv) testing to demonstrate that wall sheathing panels can supply resistance for wall bracing.

Tests of physical properties include measuring the expansion of the panels when exposed to elevated moisture conditions. Two methods for measuring linear expansion have been developed. One is the oven-dry to vacuum-pressure soak technique, which is adequate for quality control purposes, but does not reflect field conditions. Accordingly, a performance test was developed which measures linear expansion when wetting on one side for a period of 14 days. The test approximates severe exposure conditions as might occur during construction and should serve as a relative measure of linear expansion conditions that might be expected in service.

Finally, the bond durability of the panel must be demonstrated for site-built construction since panels will be called upon to resist wetting and drying and still be able to perform following exposure to these conditions. Results of extensive research have indicated that large-scale specimens exposed to hot water soaking and oven drying and then subjected to the same structural performance criteria is adequate evidence of the ability of the panel to remain intact during exposure to weathering. These tests are equivalent to several years of permanent exposure to the weather. Retention of strength following moisture cycling is also used as a measure of bond performance.

Certification of performance is a straightforward process under a performance-based system. First a sample of the candidate product is tested according to the individual performance criteria discussed above. After demonstrating adequate structural performance, physical properties, and bond durability, a span rating can then be assigned. Concurrent with performance-oriented tests, a product evaluation is also performed in which properties unique to the product and suitable for quality assurance are measured. Such properties include bending strength and stiffness, thickness, density, and others which may be useful in monitoring quality under a third-party basis. A mill specification is drawn up following this intensive product evaluation. This individualized mill specification is unique to the mill and product qualified under the standard.

2.2 *Physical Properties*

Physical property data have been collected for commercial products and this will be continuing as the industry matures.

(*a*) *Effects of moisture*. Many of the physical properties of structural panels are affected by the amount of moisture present in the wood. The EMC of structural panels is highly dependent upon RH, but is essentially independent of temperature between 0 °C and 85 °C. Examples of EMC values for structural panels at 25 °C include 4% at 40% RH, 7% at 70% RH, and 23% at 100% RH.

The expansion in length or width of structural panels from dry (less than 1% moisture content) to saturation is less than 0.5%. Linear expansion when exposed to a change in RH from 50% to 90% is typically less than 0.3%. The dimensional stability of panels exposed to liquid water also varies. Tests were conducted with panels exposed to wetting on one side as would be typical on a rain-delayed construction site. Such exposure to continuous wetting on one side of the panel for 14 days resulted in less than 0.35% expansion across or along the major axis of the panel.

(*b*) *Permeability*. The permeability of structural panels is different from solid wood in several ways. The overlapping flakes from which panels are made create many small gaps, which provide pathways for fluids to pass when entering through the panel edge. Typical permeability for untreated or uncoated ṁ in (9.5 mm) structural panels is about 0.23 g m^{-2} h^{-1} mm Hg^{-1}.

2.3 Mechanical Properties

APA–The Engineered Wood Association has collected extensive data on all types of structural-use panels. The data collection effort has been over time and across a number of mechanical properties. The results were first published in a technical note (APA 1998), and subsequently as a supplement to the consolidated industry design document (APA 1999b). The work details design principles for using structural-use panels in a variety of ways.

See also: Lumber: Laminated Veneer*; Wood: Materials for Veneers and Panels#; Wood: Nonstructural Panels*; Wood: Structural Panel Processes*

Bibliography

American National Standards Institute 1979 *Mat-formed Wood Particleboard*, ANSI A208, 1-1979. National Particleboard Association, Silver Spring, MD

American Society for Testing and Materials 1987 *Annual Book of ASTM Standards, Vol. 4.09: Wood.* ASTM, Philadelphia, PA

APA–The Engineered Wood Association 1986 *Plywood Design Specification*, Form Y510. APA, Tacoma, WA

APA–The Engineered Wood Association 1987 *Grades and Specifications*, Form J20. APA, Tacoma, WA

APA–The Engineered Wood Association 1995 *Design Capacities of APA Performance Rated Structural-use Panels*, Technical Note N375B. APA Tacoma, WA

APA–The Engineered Wood Association 1998 *Performance Rated Sidings*, Form E300. APA, Tacoma, WA

APA–The Engineered Wood Association 1999a *Allowable Stress Design for Engineered Wood Construction, Supplement: Shear Walls & Diaphragms.* AF&PA, Washington, DC

APA–The Engineered Wood Association 1999b *Allowable Stress Design for Engineered Wood Construction, Supplement: Structural-use Panels.* AF&PA, Washington, DC

Bengelsdorf M F 1981 *Linear Expansion and Thickness Swell of Wood Based Panel Products After One-side Wetting*, Report PT81-25. APA–The Engineered Wood Association, Tacoma, WA

Canadian Standards Association 1985 *Waferboard and Strandboard*, CAN 3-0437.0M85. Canadian Standards Association, Rexdale, ON, Canada

Maloney T M 1977 *Modern Particleboard and Dry Process Fiberboard Manufacturing*. Miller Freeman, San Francisco, CA

O'Halloran M R 1975 *Plywood in Hostile Environments*, Form Z820G. APA–The Engineered Wood Association, Tacoma, WA

O'Halloran M R, Youngquist J A 1983 An overview of structural panels and structural composite products. In: *Proc. Structural Wood Research State-of-the-art and Research Needs Workshop*. American Society of Civil Engineers, New York

Sellers T Jr. 1985 *Plywood and Adhesive Technology*. Dekker, New York

US Department of Commerce 1983 *Product Standard for Construction and Industrial Plywood*, PS 1. US Department of Commerce, Washington, DC

Wood A D, Johnston W, Johnston A K, Bacon G W 1963 *Plywoods of the World: Their Development, Manufacture and Application*. Morrison and Gibb, London

M. R. O'Halloran
Western Wood Products Association, Portland, Oregon, USA

Wood: Thermal Properties

Thermal mechanisms play important roles in living trees, during the conversion of wood into useful products, and in affecting the performance of such products in service. This article summarizes how the main thermal properties of wood, including its coefficients of thermal conductivity and expansion, and heat capacity, influence the way wood and wood-based composite materials may be processed and used. Thermal effects on wood formation are also summarized.

Wood consists of elaborate spatial arrangements of diverse types of cell which are themselves sophisticated rigid-walled polymeric structures with a central axial void (lumen). The cellular structure of wood renders it highly anisotropic in terms of many of its properties (see *Wood: Ultrastructure**; *Wood: Macroscopic Anatomy**): heat flow and thermal expansion differ relative to orientation within the stem. The thickness of cell walls relative to lumen size directly influences the density of wood and hence the availability of pathways for conductive heat transfer (see *Wood: Density#*). Moisture may reside in wood both in the hydrogen-bonded state in cell walls and as liquid water or water vapor in the lumens (see *Wood: Moisture Content, Hydroscopicity, and Sorption**). The form and concentration of water in wood influences all of its thermal properties.

Wood is chemically degraded by prolonged exposure to temperatures in excess of about 375 K, and charring and combustion occur at higher temperatures. The present discussion is, however, limited to the behavior of wood at temperatures below 375 K (see *Wood Products: Thermal Degradation and Fire#*). It should, however, be appreciated that the combustion of wood is dependent upon some of its intrinsic thermal properties, including conductivity and heat capacity.

1. Thermal Effects in Living Trees

Heat affects the very nature and speed of wood formation in living trees. The biochemical processes upon which the formation of xylem (wood substance) at the cambial zone (lying immediately beneath the bark) depend are highly influenced by temperature. Substantial fluctuations of temperature occur within stems both seasonally and in daily cycles. Daily fluctuations of all-important vertical temperature gradients are often dominated by the magnitude of liquid movement. Sap is drawn up from roots, via lumens of sapwood (outer annulus of the stem cross-section), to minute openings (stomata) on the underside of leaves (the process of transpiration). Significant quantities of heat are carried with the moving sap. Such movement is usually rapid during daylight hours and all but stops at night. Heat exchange at the external bark–atmosphere interface is often secondary to the effects of fluid movements within.

2. *Thermal Effects in Wood Processing*

The stiffness of wood and the deformation that it sustains under prolonged loading (see *Wood: Creep and Creep Rupture**) is co-dependent on its temperature and moisture content (wood may be termed thermo-hygro-viscoelastic). Thermally induced softening may be used to advantage during industrial procedures such as cutting and fiber separation, and in the heated consolidation of wood-based composite materials. Manipulation of temperature and moisture conditions during such processes can often be critical to their effectiveness and economic viability. In addition to temporarily softening the wood during the consolidation of composite materials in hot presses, the application of heat is also critical since adhesives used to bind the wood components together are predominantly of the thermosetting type: the rate at which they polymerize and form bonds increases almost exponentially with temperature (see *Wood: Adhesives*$^{\#}$; *Wood: Structural Panel Processes**).

Judicious application of heat is also employed to hasten the drying of lumber and the subdivided wood components used to make composite products. Drying is a necessary stage in the production of almost all wood-based products, and effecting the necessary internal heat transfer (usually while also controlling moisture movement) while minimizing damage to the structure is an important challenge.

The application of heat to wood being processed often leads to the generation of vapor pressure gradients among its pore spaces as hydrogen-bonded water in the cell walls migrates into the lumen and changes phase. Convection and diffusion of the water vapor so produced can be a very important mechanism of internal energy transfer, often exceeding that contributed by thermal conduction. Such can, for example, be the case in the kiln drying of wood and in the hot pressing of wood-based composite materials. Temperature and moisture gradient fluctuations through the thickness of a medium density fiberboard panel (see *Wood: Structural Panels**) during computer-simulated consolidation in a hot press are shown in Fig. 1. Rapid internal temperature increases are due to the combined effects of thermal conduction and the evaporation, migration, and subsequent re-condensation of water vapor. Permeability and tortuosity (respectively affecting convection and diffusion of vapor in the porous composite structure), and hygroscopicity (affecting vapor generation per unit energy input) are therefore important material properties in such situations (see *Wood: Diffusion and Permeability*$^{\#}$).

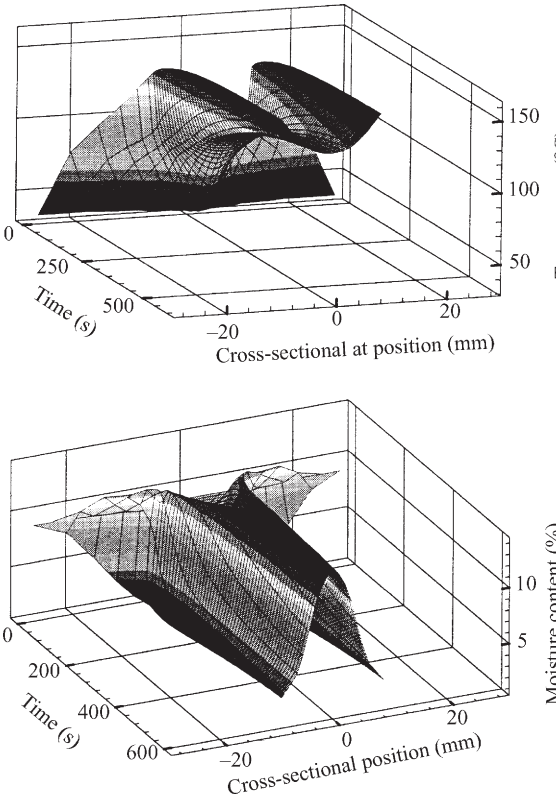


Figure 1
Temperature and moisture distributions through the thickness of a medium density fiberboard panel during consolidation in a simulated industrial hot press.

Vapor generation and convection makes measurement of the thermal properties of wood (such as thermal conductivity and heat capacity) problematic. Conventional steady-state methods in which temperature gradients are applied for prolonged periods are not appropriate. Rapid unsteady-state methods are usually necessary to avoid the onset of significant moisture migration within test specimens.

3. *Thermal Effects in Service*

Heat movement in the walls, floors, and ceilings of buildings may be minimized by taking advantage of the thermal insulating properties of wood and appropriately designed wood-based composite materials. The high volume fraction of air spaces in dried wood of low density can offer an effective barrier to heat transfer. Indeed, most thermal properties depend upon wood density since this affects the proportion of the structure that is void (occupied by air or liquid water) versus that of solid cell wall substance (see *Wood: Density*#). The density of dry wood cell wall material remains remarkably uniform among fiber types and wood species (approximately 1500 kg m^{-3}). The quantity of water adsorbed within the hygroscopic cell wall material does, however, affect thermal conductivity and consequent heat transfer rates in specific applications. Furthermore, wood structure is highly anisotropic since the wood fibers are predominantly aligned nearly vertically in the standing stem, though typically 5% are oriented radially. Such anisotropy leads to important differences in the material properties, which affect how heat and fluids move through wood and wood-based materials.

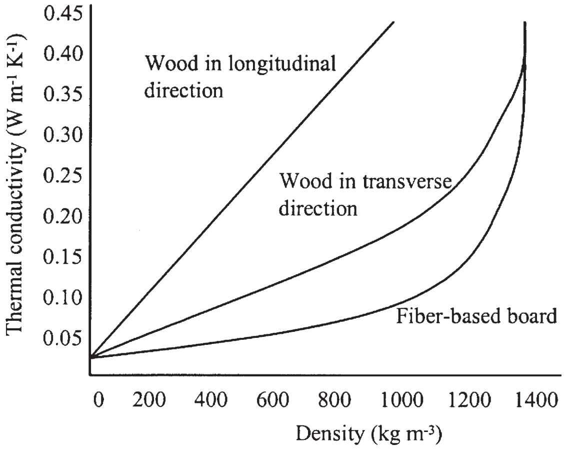


Figure 2
Effect of density on thermal conductivity of wood and two wood-based composite products at 12% moisture content and temperature of 300 K (reproduced by permission of Springer, Berlin, from *Principles of Wood Science and Technology*, Vol. 1, *Solid Wood*, 1968).

Wood undergoes only modest dimensional changes with temperature, although temperature changes of wood products in service often lead to concomitant changes of moisture content. Moisture content fluctuations below fiber saturation point (where the concentration of water molecules adsorbed within the wood cell walls varies) can cause substantial and often problematic anisotropic shrinkage and swelling or the development of internal stresses (see *Wood: Dimensional Change from Moisture*#).

4. *A Summary of Thermal Property Values*

Summarized in Table 1 are selected thermal property values of wood. Values do, however, vary widely; those readers requiring greater accuracy should refer to the literature. Properties for wood-based composites depend upon the organization of the wood constituents, adhesive, and method of consolidation, and are beyond the scope of this article. Note: moisture content values (mc) are expressed on an oven dry wood basis.

Table 1
Selected thermal property values of wood.

Thermal conductivity, κ	typical value 0.1–0.13 W m^{-1} K^{-1} (for transverse flow in a softwood of density 450 kg m^{-3} and mc of 12%)
density effect	see Fig. 2
moisture content effect	near-linear positive correlation (up to 25% mc)
flow direction effect	transverse to longitudinal ratio typically 1:1.8
Heat capacity, c	typical value 1.6 kJ kg^{-1} K^{-1} (for wood at mc of 12% and temperature of 17 °C)
density effect	minimal
temperature effect	moderate positive correlation
moisture content effect	additive for water and wood (plus a small positive adjustment factor)
Coefficient of thermal expansion	typical value 30 × 10^{-6} K^{-1} (for wood of density 450 kg m^{-3}, mc of 12%, radial direction)
density effect	minimal in longitudinal direction, near linear positive correlation in transverse
directional effect	longitudinal-to-radial ratio typically 1:7, radial-to-tangential ratio typically 1:1.4
temperature effect	small; up to 50 °C

Figure 2 shows the effect of density on k for wood and two types of composite.

Bibliography

Cermak J, Kucera J 1981 Compensation for natural temperature radiance at the measuring point during sap flow rate determination in trees. *Biologica Plantarum* **23**, 469–71

Humphrey P E 1999 The bonding speed of adhesives as a function of temperature: An automated evaluation system. In: Wolcott M *et al.* (eds.) *Proc. 33rd Int. Particleboard and Composite Materials Symp.* Washington State University, Pullman, WA

Kollman F F P, Cote W A 1968 *Principles of Wood Science and Technology*, Vol. 1, *Solid Wood.* Springer, Berlin

Thoemen H, Humphrey P E 1999 The continuous pressing process for wood-based panels: An analytical simulation model. In: Hague *et al.* (eds.) *Proc. 3rd European Panel Products Symp.* The BioComposites Centre, Bangor, UK

P. E. Humphrey
Oregon State University, Corvallis, Oregon, USA

Wood: Ultrastructure

Other articles describe wood composition and cellular structure on a macroscopic scale, but in many instances the successful processing of wood requires a more in-depth understanding of these characteristics. A brief overview of wood chemical composition is given to provide a better understanding of cellular ultrastructure.

1. Chemical Composition of Wood

Wood is composed of three primary constituents: cellulose, a linear, crystalline carbohydrate polymer with a high degree of polymerization (~10 000–14 000); hemicelluloses, a term used to collectively refer to the numerous pentosan and hexosan carbohydrate polymers having a much lower MW and a greater solvent reactivity than cellulose; and lignin, a highly cross-linked polyphenolic material.

Two so-called secondary constituents are also present in, but not an integral part of, the cell wall structure: ash (salts of minerals such as calcium, potassium, and magnesium) and extractives (any compound that is removable with a neutral solvent). Ash content is usually <0.5% for all North American species, but may be as much as 5% or more for tropical species. Silica is sometimes present in appreciable quantities. Extractives frequently add color, odor, or other characteristics such as decay resistance to wood e.g., chemicals such as oleoresins (pine "pitch") and tannins. Different species contain different types of extractives, and wood from different parts of a tree will likely contain varying amounts of extractives e.g., heartwood usually has a greater extractive content than sapwood.

Table 1 lists the amounts of cellulose, hemicelluloses, and lignin in wood compiled from the literature on North American gymnosperms (conifers, "softwoods") and angiosperms (broad-leaved trees, "hardwoods") (Browning 1963, Timell 1967, Koch 1972, 1985, Page *et al.* 1975, Panshin and de Zeeuw 1980, Weil *et al.* 1998). Q3–Q1 represents the inner quartile spread. Tropical angiosperms may have significantly more lignin than temperate-zone angiosperms (Timell 1967). A study of 20 tropical woods (mostly from Nigeria) found that lignin content ranged from 22% for iroko (*Chlorophora excelsa*) to 50% for mpingo (*Dalbergia melanoxylon*) (Sosanwo *et al.* 1995). Cellulose contents were ~35%.

There are statistically significant differences between the hemicelluloses and lignin determinations for temperate-zone angiosperms and gymnosperms, the latter having less hemicelluloses and more lignin. Variability could be from differences in analytical procedures, but it is known that trees having different genetic heritage contain differing proportions of the primary constituents, and there is also variation within trees. In gymnosperms, for example, the juvenile wood surrounding the pith of the tree has more lignin and less cellulose than wood just beneath the bark formed from a more mature cambium. There is a similar trend with tree height; topwood (with a higher juvenile wood percentage) generally has less cellulose and more lignin than bottomwood. Hemicelluloses seem to be more concentrated in the juvenile regions as well, and usually decrease by a few percent in mature wood regions (Panshin and de Zeeuw 1980).

Table 1
Primary constituents of wood (%) (moisture-, extractive- and ash-free basis).

	Cellulose	Hemicelluloses	Lignin
Gymnosperms			
Medians of 28 analyzes (Q3–Q1)	43.0 (7.7)	25.0 (11.7)	28.6 (2.9)
Minima (maxima)	38.1 (57.4)	13.3 (34.1)	23.5 (33.9)
Angiosperms			
Medians of 62 analyses (Q3–Q1)	45.8 (4.7)	31.1 (6.5)	26.4 (3.6)
Minima (maxima)	38.5 (55.2)	18.7 (39.7)	17.3 (34.7)

2. *Cell Wall Structure*

There are four essential relationships in the structure of wood cell walls: (i) aggregation of cellulose, lignin, and hemicelluloses molecules to form larger structures; (ii) arrangement of these structures to form a multilayered cell wall; (iii) variation of concentrations of cellulose, hemicelluloses, and lignin within the cell wall and the intercellular domain; and (iv) specialization of cell walls, such as interconnecting pits. Reaction wood has an abnormal cell wall structure that is described later.

2.1 Aggregation of the Primary Constituents into Larger Structures

As stated previously, cellulose is the principal constituent in cell walls. Recent work suggests that cellulose molecules are formed at macromolecular terminal complexes (TCs) in the cell plasma membrane. There is a multistage self-assembly sequence. Parallel glucan chains form planar sheets that quickly hydrogen bond to other sheets to produce mini-crystals, and associate with others to form highly-ordered cellulose microfibrils. Microfibrils are approximately square, and the atomic force microscope (AFM) has shown them to have repeat patterns along their length corresponding to the lengths of a glucose molecule and a cellobiose unit (two glucose molecules) (0.52 and 1.04 nm, respectively) (Hanley *et al.* 1992, Baker *et al.* 1998).

Measurements of bacterial cellulose and beaten kraft-pulped spruce fibers have shown that the lateral microfibril dimensions can be as small as 1.0–1.5 nm (Haigler and Weimer 1991, Okuda *et al.* 1994, Hanley 1995, Tsekos *et al.* 1996). Microfibrils have not been observed to be this small in undelignified structures. Other workers have demonstrated that the space between cellulose microfibrils is ~3–4 nm (Terashima *et al.* 1993) and that hemicelluloses at least partially cover the cellulose in native wood (Tenkanen *et al.* 1999). Hanley (1995) found that the widths of microfibrils from microtomed sections of undelignified black spruce were ~7.5 nm in diameter, probably an indication of the combined cellulose-hemicelluloses dimension. Lignin is deposited after the cell wall formation is initiated and does not affect cellulose dimensions or structure.

The wood cell wall structure is frequently thought of as a two-phase system in which cellulose is a discrete component and the lignin and hemicelluloses are blended together to form an encrusting matrix around the cellulose. Figure 1 illustrates this two-phase model, which has recently been questioned by some researchers. Studies using the cellulose-producing bacteria *Acetobacter xylinum* as a model system have shown that cellulose and hemicelluloses coaggregate to form complex celluloses. It is possible that hemicelluloses are blended with the cellulose in most higher plant cell walls (particularly in wood), and that microfibrils that are generally observed in isolated celluloses are consequences of the extraction of the hemicelluloses. Work by Salmén and Olsson

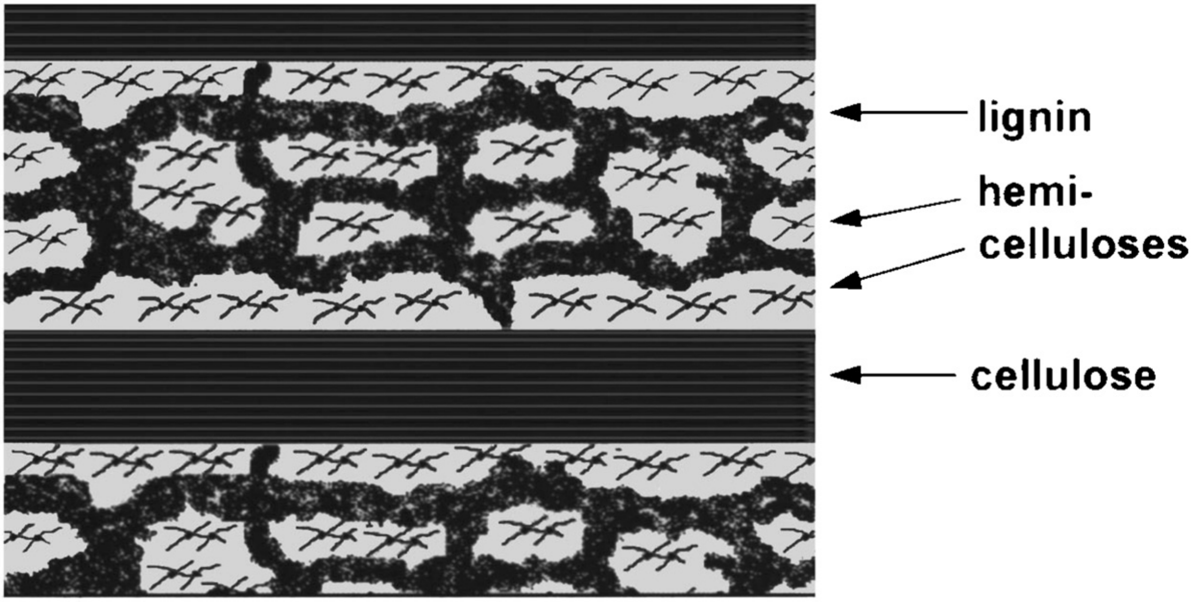


Figure 1
Two phase model for organization of the cell wall components (after Salmén and Olsson 1998).

(1998) used dynamic mechanical analysis to study spruce specimens at varying relative humidities, which indicated that hemicelluloses may be associated with both lignin and cellulose. According to their findings, xylan is associated with lignin and glucomannan with cellulose in spruce (*Picea abies*).

Molecular mechanics and dynamics studies indicate that the hemicelluloses are hydrogen bonded and oriented parallel to cellulose, and that lignin may have a helical structure (Faulon *et al.* 1994). However, the interactions are complex and many inferences remain to be verified.

X-ray diffraction studies have demonstrated departures from an infinitely extended linear lattice in native cellulose microfibrils; consequently it has often been stated that microfibrils contain both crystalline regions and less well ordered (amorphous) regions. Raman spectroscopy and solid-state ^{13}C-NMR studies have shown that this model is inadequate to describe cellulose structure. Through use of the alga *Valonia,* and other native celluloses as experimental substrates it has been shown that two different forms of cellulose (I_α and I_β) are present in the same microfibril. There is speculation that the I_α and I_β forms of cellulose may correspond to triclinic and monoclinic unit cells. In wood, the concomitant presence of these two forms probably affects x-ray crystallinity measurements to a lesser degree than cellulose chain ends or the intimate associations with hemicelluloses. There is much yet to learn about nanoscale level of organization of constituents in wood.

2.2 Multilayer Structure of Cell Walls

Wood cells are hollow layered structures with a primary cell wall on the outside and a secondary cell wall and a lumen contained within. Each layer contains oriented cellulose microfibrils together with hemicelluloses and lignin, and the microfibril orientation affects the ultimate strength of the tissue. The primary cell wall (P) is very thin and contains randomly oriented microfibrils surrounded by a lignin-rich material containing lesser amounts of cellulose, hemicelluloses, pectins, and structural glycoproteins. The secondary cell has layers visually distinguishable by differences in microfibrillar orientation; progressing from the primary cell wall inward towards the lumen, the layers are sequentially designated as S_1, S_2, and S_3. Figure 2 is a generalized diagram of a woody cell wall based on latewood tracheids of longleaf pine (*Pinus palustris*) (Dunning 1968).

The S_1 layer is slightly thicker than the primary cell wall, and the microfibrils are arranged in several thin lamellae with the outermost lamellae oriented at angles approaching 90° to the fiber axis. Two inner lamellae of the S_1 layer are arranged in alternating right- and left-hand helices at angles of $\sim 65°$ from the fiber axis, with several overlapping lamellae at progressively steeper orientations that create a gradual transition to the S_2 layer. The S_2 is the thickest layer, typically comprising around 75% of the cell wall, with the microfibrils arranged in a right-hand helix of 5–25° to the fiber axis. The S_3 layer is slightly thinner than the S_1, and similarly has several criss-crossed lamellae. The microfibril angle within the S_3 layer varies from ~ 60–$85°$, with the outermost S_3 lamellae having similar angles as the S_2, acting as transitional elements between the S_2 and S_3. Finally, most gymnosperms (and some angiosperms) contain a thin "warty layer" on the innermost lamella of the S_3, the surface of the cell lumens. The warty layer is $\leqslant 0.1\,\mu m$ thick and characterized by wart-like protuberances, which may be remnants of protoplasmic contents, and have no effect on fiber strength, although they may affect permeability.

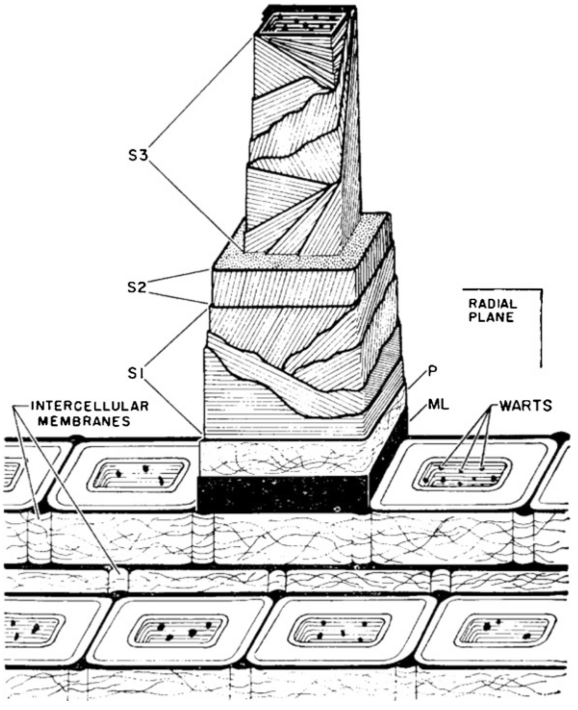


Figure 2
Cell wall structure of a longleaf pine (*P. palustris*) latewood tracheid (after Dunning 1968).

Surrounding the primary cell wall is a thin lignin-rich layer termed the middle lamella (or inner cellular layer). In early microscopic studies these layers could not be clearly distinguished and the term compound middle lamella arose to refer to the combination of the (true) middle lamella and the primary cell wall. Table 2 lists the approximate thickness of the various cell wall layers of gymnosperm tracheids. Interested readers are referred to Dunning (1968) for more detailed information and micrographs of cell wall structures.

2.3 Distribution of Chemical Constituents

The distribution of the primary chemical constituents within wood cell walls has a significant effect on wood processing operations, and investigations have been ongoing since the seminal work by Bailey (1936). Studies have relied upon a number of destructive and nondestructive techniques e.g., microdissection, acid digestion, polarizing light microscopy, UV-microspectroscopy, enzyme-gold complex studies, Raman spectroscopy, bromine-tagged lignin measurements using x-ray spectroscopy, etc. All of these techniques have some limitations, and some contradictory results have been published that are difficult to resolve. General agreement exists on the qualitative distribution of cellulose, hemicelluloses, and lignin in the middle lamella and the secondary cell wall, but there is considerable disagreement on quantitative values. Table 3 and Fig. 3 show one possible distribution of cellulose, hemicelluloses, and lignin within the various cell wall layers of gymnosperms based on reports from Meier (1961), Fergus *et al.* (1969), Wood and Goring (1971), Saka *et al.* (1978), Hardell and Westermark (1981), Donaldson (1987), and Westermark *et al.* (1988).

When the nominal cell wall layer thicknesses listed in Table 2 are used to calculate the overall chemical composition, the results compare favorably to those presented in Table 1 (i.e., about 42% cellulose compared to 43% in Table 1, 32% hemicelluloses vs. 25%, and 25% lignin vs. 29%). In spite of the greater concentration of lignin in the compound middle lamella, it must be apparent that most of the lignin is actually present within the secondary cell wall due to its greater thickness.

Table 2
Cell wall layer thicknesses of gymnosperm tracheids (μm)

Middle lamella	⩾0.1
Primary cell wall	0.03
S_1	0.2
S_2	⩾5
S_3	0.07–0.08
Warty layer	⩽0.1

Source: Kollmann and Côté (1968)

2.4 Cell Walls and Distinguishing Features

Up to this point the woody cell wall has been described in terms of chemical composition and structural organization of the principal constituents in the cell wall. However, wood cells are not balloon-like bodies with smooth walls and closed ends, but have many features that characterize the cell walls. Foremost are the paired openings in adjacent cell walls that permit the passage of fluids. In both gymnosperms and angiosperms some cell walls are penetrated by openings called pits. Angiosperms also have openings (perforation plates) at the junctures between ends of individual vessel elements. Other features are less uniformly present and are of diagnostic significance to identify wood species.

(*a*) *Gymnosperm pits.* In gymnosperms pits are found in three locations: intersections of longitudinal parenchyma and ray parenchyma cells, radial walls of adjacent longitudinal tracheids, and intersections of longitudinal tracheids with ray parenchyma or ray tracheids. Simple pits between adjacent parenchyma cells are thin, membrane-like structures comprising

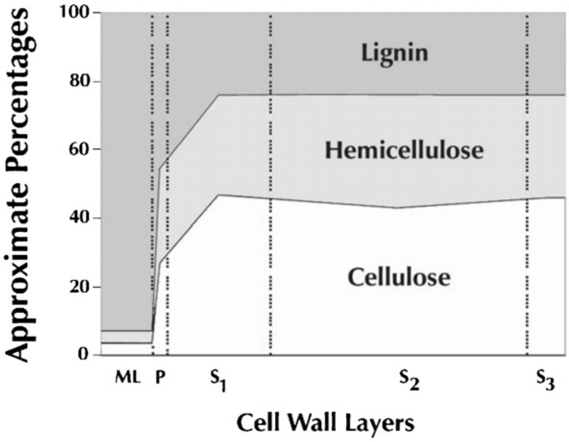


Figure 3
Possible cellulose, lignin and hemicellulose distribution in a gymnosperm cell wall.

Table 3
Gymnosperm cell wall layer constituent (%).

Cell wall layer	Cellulose	Hemicelluloses	Lignin
Middle lamella	3.5	3.5	93
Primary cell wall	27	27	46
Compound middle lamella	10	10	80
S_1	47	29	24
S_2	43	33	24
S_3	46	30	24

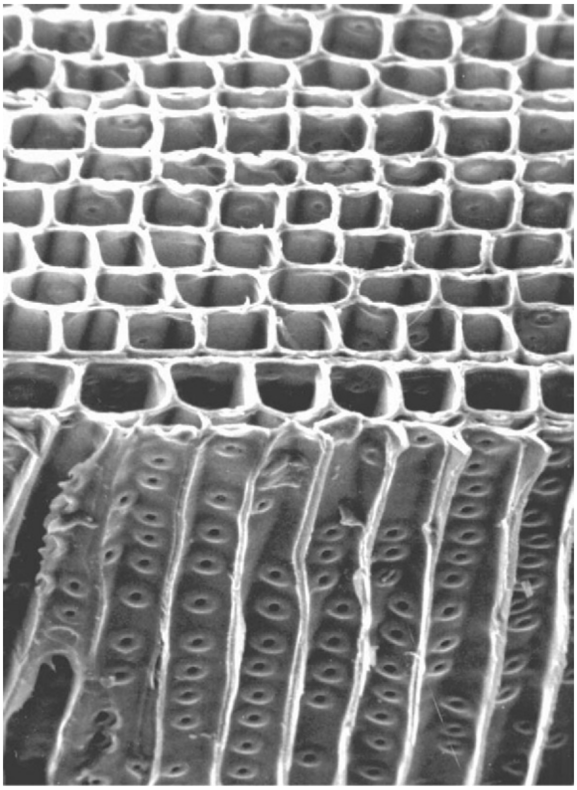


Figure 4
Longitudinal tracheids showing extensive bordered pitting on the radial cell walls of Jack pine (*P. banksiana*) (courtesy of the Institute of Paper Science and Technology).

the middle lamella and the primary cell walls. A bordered pit pair has overarching structures on the lumen side of each cell wall surrounding apertures between two adjacent longitudinal tracheids (Fig. 4). Each bordered pit pair encloses a web-like microfibril structure called a *margo* that supports (except in western redcedar (*Thuja plicata*)) a central and much less permeable *torus*, the assembly acting as a valve to open or close the openings between tracheids (Fig. 5). When the torus is moved enough to block the pit aperture, as by surface tension during wood drying, the pit is said to be aspirated (Fig. 6). Longitudinal tracheids and ray tracheids are also interconnected via bordered pits, but the pits are much smaller than those between longitudinal tracheids.

Pits occurring at the crossfields between ray parenchyma and longitudinal tracheids are termed half-bordered pits because they have the characteristics of bordered pits on the longitudinal tracheid and simple pits on the ray parenchyma cell. These pits can be classified into five categories by differences in the shapes of apertures and overhanging borders: fenestriform (also called windowlike), pinoid, piceoid, taxodioid, and cupressoid. Figure 7 illustrates a fenestriform crossfield pit in eastern white pine (*P. strobus*). The ray is heterocellular, and the cell types are thereby distinguishable. For more information and examples of crossfield pits the reader is referred to Panshin and de Zeeuw (1980).

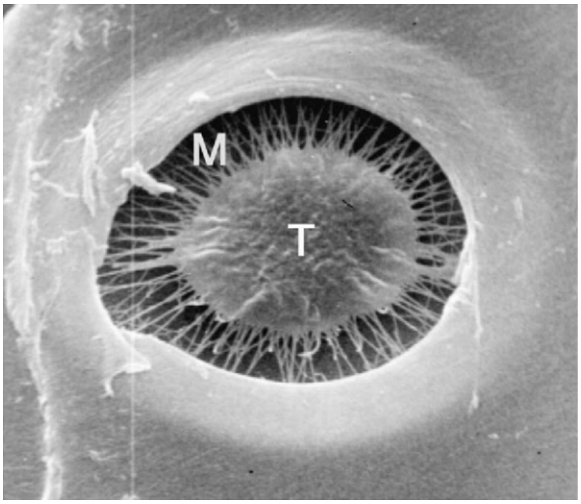


Figure 5
Margo (M) and torus (T) from an earlywood tracheid of balsam fir (*Abies balsamea*); the lumen-side pit border has been removed in this micrograph (courtesy of the Institute of Paper Science and Technology).

(*b*) *Angiosperm pits*. Since numerous cell types may be present in angiosperms, pitting can be complex, including simple, half-bordered, and full-bordered types. Except for pitting between adjacent vessel elements, the pit appearance on the various types of cells has little diagnostic value. Intervessel pitting (largely on radial walls) is designated as opposite or alternate. Opposite pitting is arranged in more-or-less regular horizontal rows like a checkerboard while alternate pitting is either unordered or arranged somewhat diagonally. A third type of pitting, perforation plates, is found at adjoining ends of vessel elements, having simple openings or more complex structures. In commercial temperate-zone species only two types of perforation plates are found: simple and scalariform. Simple perforation plates are uncomplicated openings surrounded by a thin rim on the ends of the vessel elements, however, scalariform are characterized by numbers of bar-like portions of the cell wall that extend across the vessel openings (Fig. 8). The type of opening and the number of bars are most readily seen on the radial surface and have diagnostic significance.

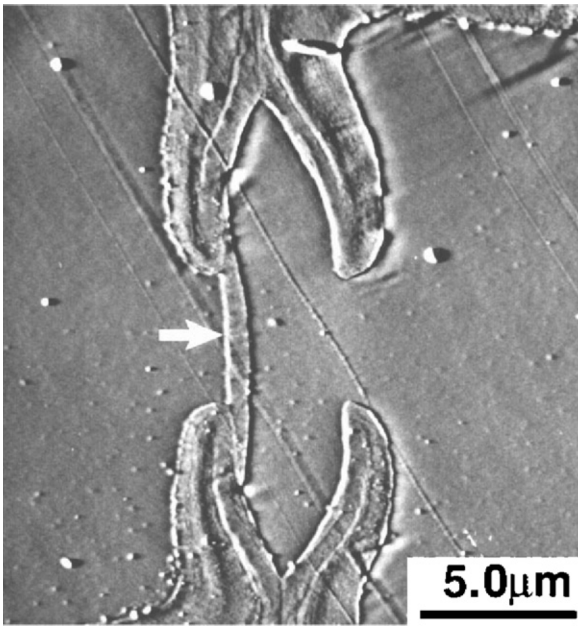


Figure 6
Cross-section of a bordered pit pair. Arrow indicates the aspirated torus blocking one of the pit apertures (after Hanley 1995).

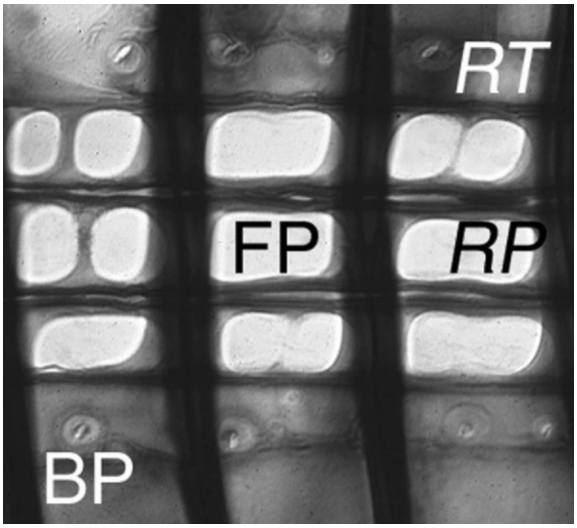


Figure 7
This radial view of a piece of eastern white pine contains both ray parenchyma (*RP*) and ray tracheid (*RT*). FP = fenestriform-type of crossfield pits, and BP = the small bordered pits between the ray tracheids and longitudinal tracheids.

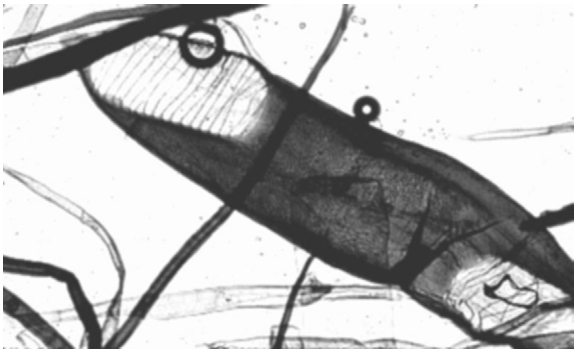


Figure 8
Birch (*Betula* species) vessel element with scalariform perforation plates.

(*c*) *Cell wall sculpturing*. Wood cell walls may be ornately structured, termed sculpturing. The most frequently observed marking is helical (also called spiral) thickening, which is a coil-like pattern of microfibrillar ridges on the lumen surface. These are a species-specific trait and can be found in gymnosperm tracheids, and in angiosperm vessel elements and vascular tracheids (and occasionally fibers). The feature is distinctive and easily observed using light microscopy. A good example of a North American coniferous species containing this structure is Douglas fir (*Pseudotsuga menziesii*) (Fig. 9).

(*d*) *Inclusions*. Two types of inclusions are commonly found in lumens of wood cells: mineral crystals and tyloses. Crystals are found in a limited number of species (e.g., iroko (*C. excelsa*), teak (*Tectona grandis*), red mulberry (*Morus rubra*)) and may even be visible to the naked eye (Fig. 10). Tyloses are balloon-like membrane structures that grow into vessel elements from adjacent parenchyma cells (Fig. 11). Tyloses are more common in some species than others; in species where they are usually absent, they may be found in heartwood. Tyloses can block the longitudinal passage of air or other fluids, and may affect the selection of wood for certain uses. Tight cooperage, for example, makes use of white oak (*Quercus species*) with its plentiful tyloses.

3. Reaction Wood

Reaction wood is the generic term used to refer to wood that has atypical anatomical and chemical characteristics, usually from its origin in leaning stems or branches. In gymnosperms this wood is called compression wood and in angiosperms, tension wood. Each form of reaction wood is distinctly different from normal cell tissues.

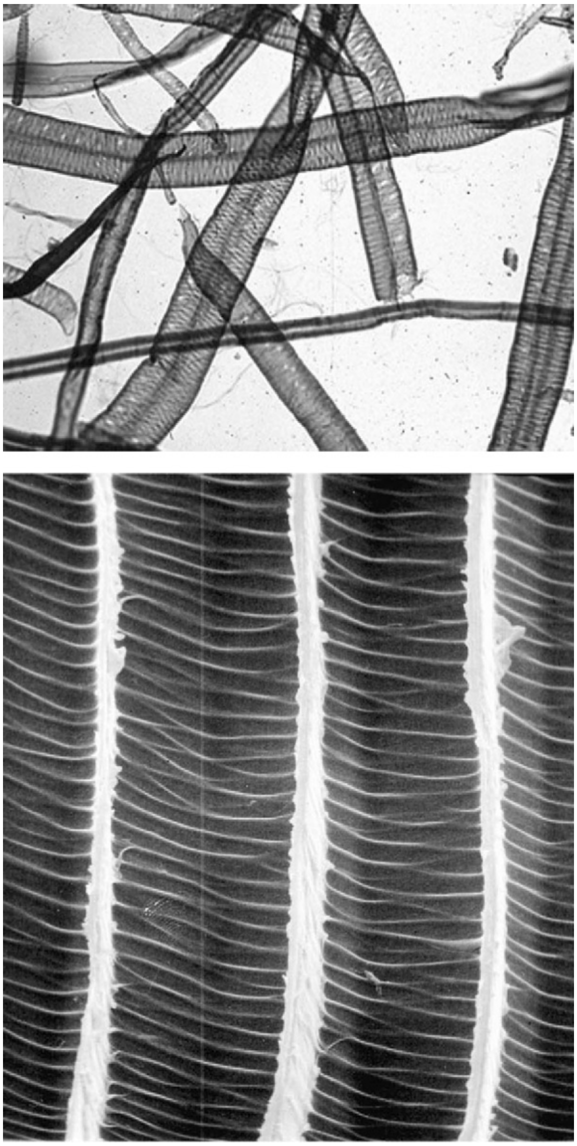


Figure 9
Transmitted light and scanning electron micrographs of Douglas-fir (*P. menzisii*) tracheids showing spiral thickenings.

3.1 Compression Wood

Compression wood is found on the underside of leaning stems and branches of conifers. Its cells are characterized by unusually high lignin and lower cellulose content. The cell walls contain spiral gaps (called "spiral checks") along the steeper S_2 microfibril angle that is a characteristic of these tissues (up to $\sim 45^\circ$ instead of the normal 5–25° from the tracheid axis) (Fig. 12). The cell walls of longitudinal

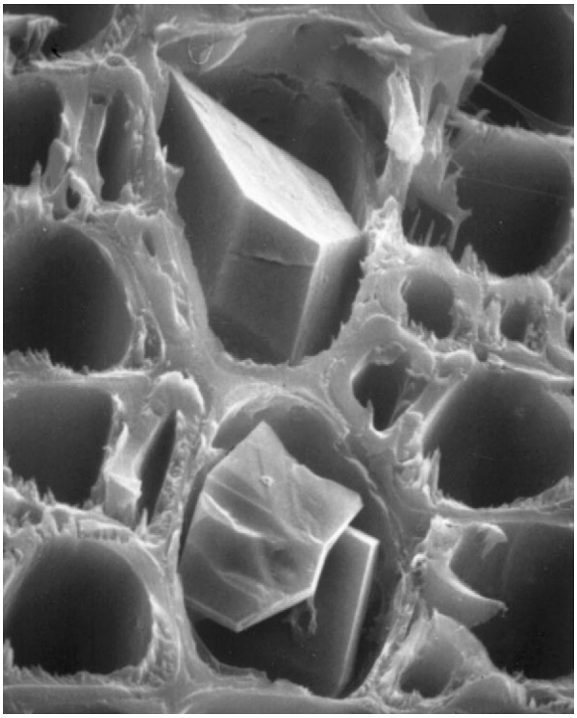


Figure 10
Mineral crystals in iroko (*C. excelsa*) (courtesy of the Institute of Paper Science and Technology).

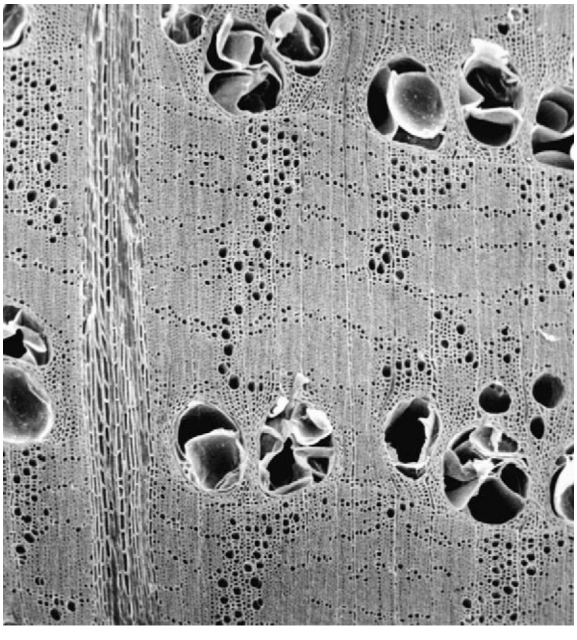


Figure 11
Tyloses in vessel elements of white oak (*Q. alba*) (courtesy of the Institute of Paper Science and Technology).

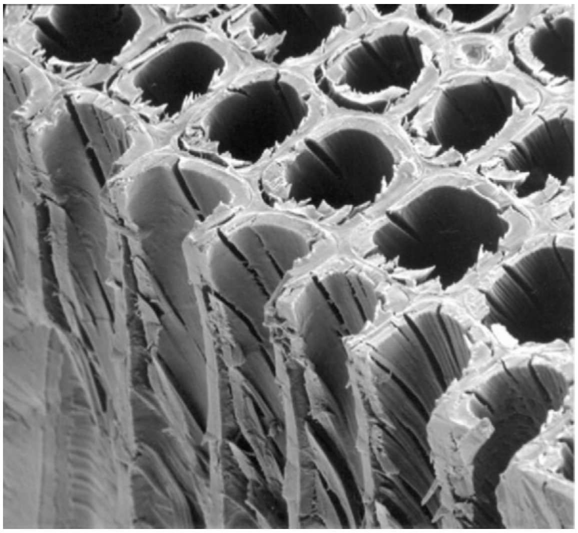


Figure 12
Compression wood in a piece of southern pine showing rounded longitudinal tracheids and spiral checking (micrograph by Kuklinsky R F, Mississippi State University Electron Microscopy Center).

tracheids are unusually rounded compared to normal tracheids, and there are intercellular spaces at some cell corners. The cell walls are frequently slightly thicker than normal wood and the S_3 layer is typically lacking.

3.2 Tension Wood

Tension wood is found on the upper side of leaning stems and branches in angiosperms. It contains more cellulose and less lignin, and the fiber structure is markedly different from that of normal wood. Tension wood fibers are characterized by a layer of microfibrils that is almost pure cellulose, the gelatinous or G-layer, and the fibers are known as gelatinous fibers. (Gelatinous fibers may also be found intermixed with normal fibers in oaks even when no reaction wood is present.) The G-layer occurs on the lumen side of the fibers and is easily detached from the rest of the cell wall (Fig. 13), causing a distinctive fuzzy appearance on the surface of machined wood.

Gelatinous fibers can be found in several types of cell wall configurations. The G-layer may occur in addition to the more normal S_1–S_2–S_3 configuration, replace the S_3 layer, or occur in place of both S_2 and S_3. The cellulose microfibrils in the G-layer are oriented at very small angles to the longitudinal fiber axis.

4. Concluding Remarks

The ultrastructure of wood is a very useful diagnostic tool for species identification and also explains some of the gross properties of wood. However, despite the use of many analytical techniques, details of the molecular ultrastructure of wood remain to be resolved.

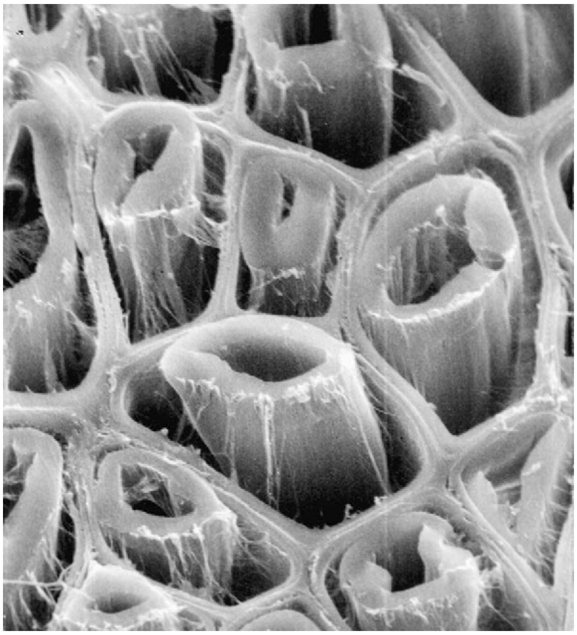


Figure 13
Tension wood showing the G-layer detached from the rest of the cell wall (courtesy of the Institute of Paper Science and Technology).

5. Acknowledgments

The author extends his sincere appreciation to the various individuals and institutions that contributed granted permission to reproduce some of the illustrations in this chapter. These include The Institute of Paper Science and Technology (formerly the Institute of Paper Chemistry), Charles Dunning, Shaune J. Hanley, Derek Gray, and Lennart Salmén. Thanks are also due to the Journal of Pulp and Paper Science, the Mississippi State University Electron Microscopy Center and to Rajai Atalla for reading a portion of an early draft about cellulose structure.

See also: Wood, Constituents of*; Wood: Diffusion and Permeability#; Wood: Macroscopic Anatomy*; Wood Products: Weathering*

Bibliography

Bailey A J 1936 Lignin in Douglas-fir—Composition of the middle lamella. *Ind. Eng. Chem., Anal. Ed.* **8**, pp. 52–5

Baker A A, Helbert W, Sugiyama J, Miles M J 1998 Surface structure of native cellulose microcrystals by AFM. *Appl. Phys. A: Mater. Sci. Process.* **66** (S), S559–63

Browning B L (ed.) 1963 *The Chemistry of Wood*. Interscience, New York, p. 689

Core H A, Côté W A, Day A C 1976 *Wood Structure and Identification*, Syracuse Wood Science Series 6. Syracuse University Press, Syracuse, NY, p. 168

Côté W A (ed.) 1965 *Cellular Ultrastructure of Woody Plants*. Syracuse University Press, Syracuse, NY, p. 603

Donaldson L A 1987 S3 lignin concentration in radiata pine tracheids. *Wood Sci. Technol.* **21** (3), 227–34

Dunning C E 1968 An investigation of longleaf pine cell-wall morphology by electron microscopy of single fibers. Ph.D. thesis, Institute of Paper Chemistry, Appleton, WI, p. 409

Faulon J L, Carlson G A, Hatcher P G 1994 A 3-dimensional model for lignocellulose from gymnospermous wood. *Org. Geochem.* **21** (12), 1169–79

Fergus B J, Procter A R, Scott J A N, Goring D A I 1969 The distribution of lignin in sprucewood as determined by ultraviolet microscopy. *Wood Sci. Technol.* **3**, 117–38

Haigler C H, Weimer P J (eds.) 1991 *Biosynthesis and Biodegradation of Cellulose*. Dekker, New York, p. 710

Hanley S J 1995 Application of atomic force microscopy to cellulose, wood, Kraft fibres and paper. Ph.D. thesis, McGill University, Montreal, Canada, p. 129

Hanley S J, Giasson J, Revol J F, Gray D G 1992 Atomic force microscopy of cellulose microfibrils—Comparison with transmission electron microscopy. *Polymer* **33** (21), 4639–42

Hardell H, Westermark U 1981 In: *Ekman-Days Int. Symp. Wood and Pulping Chemistry*, Stockholm. **5**, pp.17–19

Koch P 1972 Utilization of the Southern Pines, Vol. 1, Agriculture Handbook No. 420. US Government Printing Office, Washington, DC, p. 734

Koch P 1985 Utilization of Hardwoods Growing on Southern Pine Sites, Vol. 1. Agriculture Handbook No. 605. US Government Printing Office, Washington, DC, p. 1418

Kollmann F F P, Côté W A 1968 *Principles of Wood Science and Technology*. Springer, New York p. 592

Meier H 1961 The distribution of polysaccharides in wood fibers. *J. Polym. Sci.* **51**, 11–8

Okuda K, Tsekos I, Brown R M 1994 Cellulose microfibril assembly in *Erythrocladia subintegra* Rosenv.—an ideal system for understanding the relationship between synthesizing complexes (TCs) and microfibril crystallization. *Protoplasma* **180** (1–2), 49–58

Page D H, El-Hosseiny F, Bidmade M L, Binet R 1975 Birefringence and the Chemical Composition of Wood Pulp Fibres. Pulp and Paper Reports PPR/137. Pulp and Paper Research Institute of Canada, Pointe Claire, Quebec, Canada

Panshin A J, de Zeeuw C 1980 *Textbook of Wood Technology*, 4th edn. McGraw-Hill, New York p. 722

Pinto B M (ed.) 1999 *Comprehensive Natural Products Chemistry*. Elsevier, Oxford, Vol. 3

Saka S, Thomas R J, Gratzl J S 1978 Lignin distribution: Determination by energy-dispersive analysis of x-rays. *Tappi* **61** (1), 73–6

Salmén L, Olsson A-M 1998 Interaction between hemicellulosess, lignin and cellulose: Structure-property relationships. *J. Pulp Paper Sci.* **24** (3), 99–103

Sosanwo O A, Fawcett A H, Apperly D 1995 C-13 CP-MAS NMR Spectra of tropical hardwoods. *Polym. Int.* **36** (3), 247–59

Tenkanen M, Tamminen T, Hortling B 1999 Investigation of lignin-carbohydrate complexes in kraft pulps by selective enzymatic treatments. *Appl. Microbiol. Biotechnol.* **51** (2), 241–8

Terashima N, Fukushima K, He L-F, Takabe K 1993 Comprehensive Model of the Lignified Plant Cell Wall. In: Int. Symp. Forage Cell Wall Structure and Digestibility, 1991. American Society of Agronomy-Crop Science, Madison, Wisconsin, pp. 247–70

Timell T E 1967 Recent progress in the chemistry of wood hemicellulosess. *Wood Sci. Technol.* **1**, 45–70

Tsekos I, Okuda K, Brown R M 1996 The formation and development of cellulose-synthesizing linear terminal complexes (TCs) in the plasma membrane of the marine red alga *Erythrocladia subintegra* Rosenv. *Protoplasma* **193** (1–4), 33–45

Weil J, Brewer M, Hendrickson R, Sarikaya A, Ladisch M R 1998 Continuous pH monitoring during pretreatment of yellow-poplar wood sawdust by pressure cooking in water. *Appl. Biochem. Biotechnol.* **70** (2), 99–111

Westermark U, Lidbrandt O, Eriksson I 1988 Lignin distribution in spruce (*P. abies*) determined by mercurization with SEM-EDXA technique. *Wood Sci. Technol.* **22**, 243–50

Wood J R, Goring D A I 1971 The distribution of lignin in stem wood and branch wood of Douglas fir. Technical Paper T95. *Pulp Pap. Mag. Can.* **72** (3), 61–8

T. E. Conners

Forest Products Laboratory, Mississippi, USA

Wood–Plastic Composites

Wood has commonly been used as inexpensive filler by the plastics industry. During the late 1980s, researchers and industries began investigating high levels of filler content (40–70%) and coupling agents to encourage interaction between the wood and thermoplastic components. An improved sophistication in processing and formulations led to the development of wood-plastic composites (WPCs) that exhibit synergistic material properties.

WPCs consist primarily of wood and thermoplastic polymers combined using technologies similar to other filled thermoplastics. While often seeking improved strength, stiffness, creep, and thermal stability, the commercial success of these materials has been primarily driven by improved moisture performance, recycled and waste material utilization, and efficiency in product and process design.

1. *Products and Applications*

WPCs have been used in many of the molded applications common to commodity thermoplastics. However, the improved mechanical, thermal, and processing characteristics of the materials have facilitated extensive use in highly contoured parts for the automotive and building products markets (Laver 1996, Puppin 1999). Interior automotive components experience significant thermal changes that provide performance challenges to traditional thermoplastics.

The increased thermal stability imparted by the addition of 40–50% wood improves the product performance. The thermoplastic component imparts improved levels of moisture performance and thermal formability to the wood. The low impact strength of WPCs does limit their use to nonstructural components like door, roof, and trunk liners.

The most rapidly growing application for WPCs is in extruded sections for residential building markets. Poly(vinyl chloride)–wood compounds are becoming common for replacement window lineals (Fig. 1). A solid thermoplastic cap can bond with the WPC to control color stability while the WPC core facilitates an easily processed, high modulus section. Low- and high-density poly(ethylene) polymers are compounded with up to 70% mass fraction wood to produce deck boards and railing. The increased thermal stability of the WPC facilitates the production of detailed sections that can be engineered for load-bearing applications (Fig. 2).

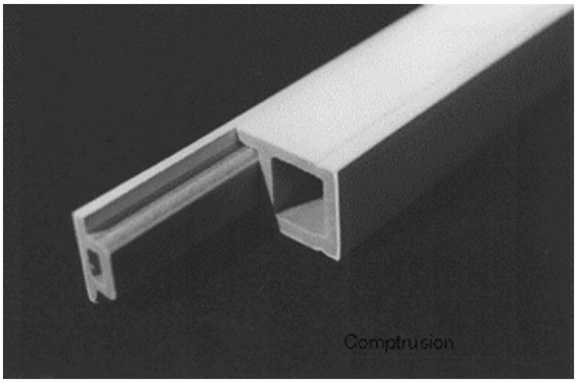


Figure 1
Wood–PVC window component capped with PVC (courtesy of Strandex Corp.).

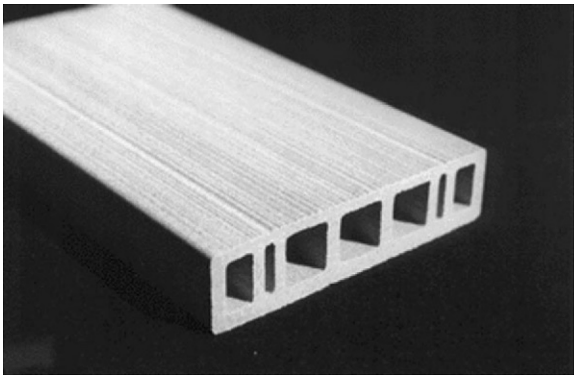


Figure 2
Hollow wood–HDPE deck board (courtesy of Strandex Corp.).

2. *Raw Materials*

WPC formulations are comprised primarily of a thermoplastic and lignocellulosic component. However, a number of additional components are commonly used to improve both the processing and properties of the resulting materials.

Wood is the most common source of the lignocellulosic component of WPCs. A variety of species is used and where species differences may exist, they have not been explicitly studied. Many agricultural fibers have also been successfully used in WPCs. Agricultural fibers often display certain advantages with respect to aspect ratio and specific mechanical properties (Sanadi *et al.* 1997, Tjong *et al.* 1999). For both wood and agricultural fibers, the material is used in fiber, flour, or particle form (Wolcott and Englund 1999). Particles and flour are formed using hammermills and grinders common to wood-based particleboard. Fibers can be produced with thermomechanical and chemical pulping methods that are used for pulp and fiberboard manufacture.

The primary criteria in selecting thermoplastics for WPCs are those with a processing temperature less than ca. 250 °C. Above this threshold, unmodified lignocellulosic materials undergo rapid thermal degradation. This thermal criterion has limited WPCs to polyolefins. Commonly used thermoplastics include poly(propylene) (PP), high and low density poly(ethylene) (HDPE, LDPE), poly(vinyl chloride) (PVC), and poly(styrene) (PS). In addition, impact modified versions of PP and PS have been explored. Both recycled and virgin sources of these commodity thermoplastics are commonly employed in powder or pellet forms.

Several minor components are added to WPCs to assist in performance of the material during processing or end use. Most commercial formulations for extrusion include lubricants that can be in the form of waxes or other solid components (Laver 1996). These lubricants primarily facilitate flow in the extruder barrels but may impede moisture movement in the product. Compounds composed of PVC typically utilize various stabilizers and plasticizers that are common to most PVC compounds (Zehner 2000). These components mainly act to alter the PVC component but the polar plasticizer components could conceivably alter the wood polymers.

3. *Wood and Thermoplastic Compatibility and Dispersion*

The most varied of components added to WPC formulations are materials aimed at improving the compatibility of the wood and thermoplastic. These components include surfactants, copolymers, and reactive compounds (Lu *et al.* 2000). The energetic differences that exist between the wood and thermoplastic impede both adhesion and dispersion of the two phases. The chemical treatments may be applied prior to composite processing. Preliminary applications typically include coating or copolymerization in solvents. However, commercial operations most commonly use methods of direct mixing in dry blending, melt compounding, or extrusion steps. In many of these compounds both the chemical structure and the modification procedure have great influence on the effectiveness of a coupling agent.

The most popular coupling agents include maleic anhydride (MA), maleic anhydride modified poly(propylene) (MAPP), isocyanates, and silanes. MA, MAPP, and PMPPIC (poly(methylene) poly(phenyl) isocyanate) are very effective for improving the mechanical properties of composites presumably by improving the wood-thermoplastic interface. Silanes exhibit some coupling effects but seem more dependent on the types of plastic. Some physical methods like corona and plasma discharges can have a positive influence on mechanical properties of WPCs (Young *et al.* 1995). Most coupling agents are proposed to facilitate chemical linkages and/or hydrogen bonds between the wood and thermoplastic. However, the wood-thermoplastic adhesion is believed to be dominated by physical adsorption and acid/base interaction if any.

Although the interaction between unmodified wood and thermoplastics is limited at best, commercial materials are often produced with no chemical treatment. Physical means of improving component dispersion are most commonly employed. Both high energy mixing methods and decreased particle sizing have been shown to increase dispersion of wood within thermoplastics (Wolcott and Englund 1999). Compression-molded composites without prior compounding tends to produce structures with stacked wood elements leading to poor moisture performance (Fig. 3). In addition, penetration of the thermoplastic into the void structure of the wood has been observed (Fig. 4). Small particle sizes increase dispersion and mechanical properties when coupling agents are not employed. This fact bolsters the use of wood flour in commercial applications that do not employ chemical treatments.

4. Physical and Mechanical Properties

The combination of wood and thermoplastics results in materials with unique physical properties. In general, the wood component tends to increase the stiffness, thermal stability, UV resistance, and workability of the composite. Thermoplastics improve moisture resistance, resistance to decay, and impart thermoforming characteristics. Mechanical properties are often the most studied of the physical properties, although other characteristics like thermal and moisture stability seem to drive commercialization. In

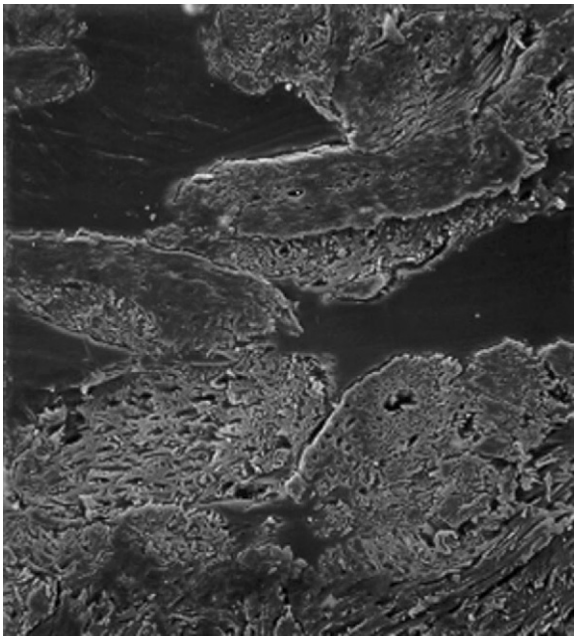


Figure 3
Compression molded wood–PS composite with wood networks that improve mechanical performance and increase moisture absorption.

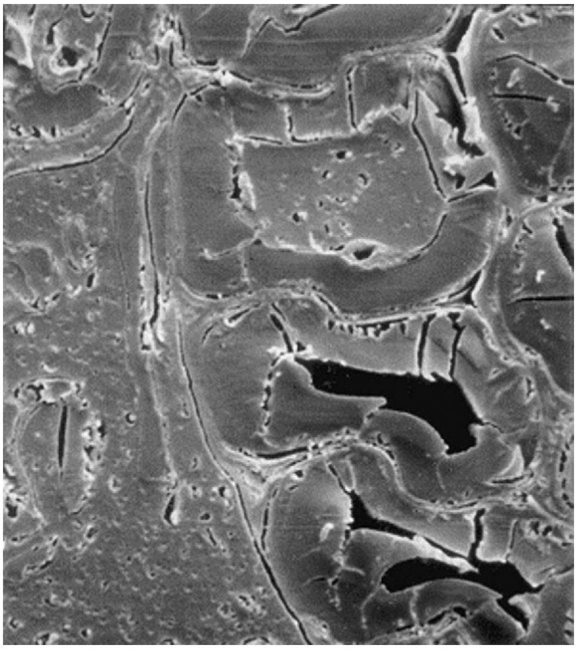


Figure 4
Extruded wood–PVC composite displaying processing induced wood damage, plastic filled lumens, and interfacial gaps.

Table 1
Tensile properties of TWP composites using wood fiber and flour with PP, PS, HDPE, and LDPE.

		Wood fiber content (wt.%)					
Property	Polymer	0	10–15	20–25	30–35	40	50
Modulus (GPa)	PP	0.9–1.1	1.1–1.22	1.2–1.55	1.2–2.0	2.4	1.2–2.85
	PS	1.5–1.9	1.7–.0	1.8–2.5	1.9–2.9	2.0–3.3	3.95
	HDPE	0.81–1.2	0.8–1.1	1.1–1.5	1.1–1.8	1.3–1.8	1.55–2.5
	LDPE	0.33	0.43	0.63	0.72	0.77	
Strength (MPa)	PP	27–34	23.5	23–25	16–21	20	11.9–27
	PS	40–42	35–36	35–36	33–34	29.9	
	HDPE	25–29	24–26.5	19–25	18–27	15–20	19
	LDPE	10.8	13.1	12.2	11.2	10.7	

Source: Wolcott and Englund (1999).

all cases, the material performance is significantly influenced by the processing methods used.

Mechanical properties of WPCs are influenced by composition in different ways. Elastic modulus increases with increasing wood fractions. Although the trend is consistent across the type of thermoplastic and wood particle (Table 1), the magnitude of the increase appears to be influenced by several factors. Composite strength is highly dependent on several factors, including the wood form, plastic type, processing method, and wood–thermoplastic interaction. In unmodified composite systems, small particle sizes are more effective at increasing the elastic modulus, favoring the use of wood flour (Törmälä *et al.* 1985). However, it appears more difficult to improve strength properties with wood flour. Decreased strength properties are generally cited with increasing wood content with a few exceptions. Woodhams *et al.* (1984, 1993) found that the tensile strength of wood fiber/PP WPCs produced with a dispersing agent or high shear mixer increase with increasing wood content (Fig. 5). Impact strength decreases with fiber loading in all cases, although elastomeric coupling agents lessen the effect (Oksman and Clemons 1998).

As with many reinforced composites, impact energy decreases with increased fiber content for wood flour and fibers (Oksman and Clemons 1998). Methods used to address this deficiency include toughened thermoplastic or matrix/fiber interphase region, optimized particle geometry, and foaming of the composite (Matuana *et al.* 1998). Introduction of a controlled cellular structure in both modified and unmodified WPCs can significantly increase the impact energy of a composite. However, the strength and stiffness typically decreases in proportion to the density. Polar elastomeric copolymers can increase impact energy while maintaining other properties, especially in maleated versions.

Moisture absorption and swelling performance of WPCs improve with increasing thermoplastic fraction. The wood moisture content during processing is generally below 1% and therefore will absorb moisture when exposed. Whereas equilibrium moisture contents well below 3% are noted with exposure to water vapor, liquid water soaks can result in moisture uptake greater than 10–20% depending on formulation. Researchers have reported that water absorption and swelling decreased disproportionately to the amount of plastic added (Wolcott and Englund 1999). These findings may indicate that material structure changes with thermoplastic content have a strong influence on material properties.

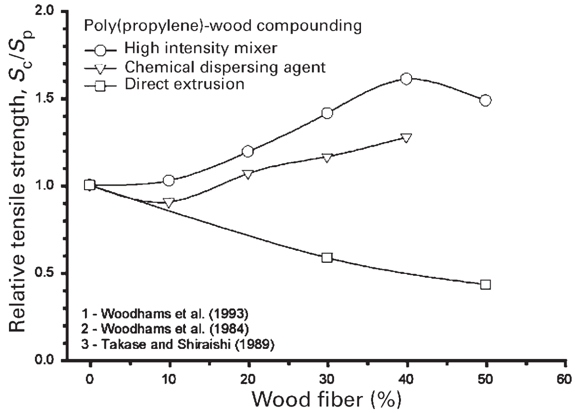


Figure 5
Tensile strength of wood-PP composites. Increased wood content can improve composite strength properties with mechanically or chemically facilitated dispersion (after Wolcott and Englund 1999).

5. Melt Blending and Compounding of Components

Dispersion of wood throughout the thermoplastic melt is a major processing step that influences the quality of the finished product. Melt blending wood and thermoplastics is facilitated through input of high shear forces. Wood flour is generally easy to disperse in a matrix but fibers and larger particles require longer times and higher mixing intensity. Twin-screw extruders provide simultaneous mixing with extrusion or injection molding. They are often used in conjunction with dry blending. Intensive-shear mixers can achieve preparation of wood fiber and thermoplastics before injection molding (Park and Balatinecz 1997).

The blending conditions play an important role in the final parameters used in processing and properties of the resulting product. Processing temperature, filler type and percentage, mixing time, and rotation speed are important parameters that govern the blending procedure. Blend properties such as viscosity, melt elasticity, and fiber length are all influenced by the mixing procedures used. Higher temperature decreases polymer viscosity and can facilitate mixing. However, thermal decomposition of lignocellulosics can become significant in the processing range of polypropylene (180–200 °C). Effectiveness can often be maximized at intermediate temperatures.

Filler dispersion in the thermoplastic matrix can depend on the duration and speed of mixing. However, the optimal time and rotation rate of mixing are controlled by the polymer type, filler properties, and mixing equipment. Both long compounding times and high rotation rates will independently increase the dispersion of fillers throughout the matrix. However, these same conditions can degrade or pulverize fibers. Takase and Shiraishi (1989) noted that tensile strength and stiffness of the composite reaches a maximum value after a finite time interval. However, composite ductility continued to increase with increased mixing time and rotation rates. The extent of fiber damage during mixing depends on fiber and polymer type, processing or mixing method, and fiber–fiber interaction.

6. Conclusions

WPCs are emerging hybrid materials that combine lignocellulosics and synthetic thermoplastics. As with synthetic polymer materials, this thermoplastic version of wood composites adds several processing advantages. Materials can be thermoformed using compression molding of sheet products, extrusion, or injection molding. The thermoforming capabilities also facilitate recycling capabilities for the product. Most WPC products are being used in exterior applications where water resistance is paramount. However, control of product shape through processing techniques may be of equal value in the end.

See also: Polymer Modified Wood*; Wood Products: Thermal Degradation and Fire#; Wood: Surface Properties and Adhesion#

Bibliography

Laver T C 1996 Extruded synthetic wood composition and method for making same. *US Patent 5516472*

Lu J Z, Wu Q, McNabbs H S 2000 Chemical coupling in wood fiber and polymer composites: A review of coupling agents and treatments. *Wood & Fiber Sci.* **32** (1), 88–104

Matuana L M, Park C B, Balatinecz J J 1998 Cell morphology and property relationships of microcellular foamed PVC/wood-fiber composites. *Polym. Eng. Sci.* **11**, 1862–72

Oksman K, Clemons C 1998 Mechanical properties and morphology of impact modified polypropylene–wood flour composites. *J. Appl. Polym. Sci.* **67**, 1503–13

Park B D, Balatinecz J J 1997 A comparison of compounding processes for wood-fiber/thermoplastic composites. *Polym. Compos.* **18**, 425–31

Puppin G 1999 Resin and wood fiber composite profile extrusion method. *US Patent 5882564*

Sanadi A R, Caulfield D F, Jacobson R E 1997 Agro-fiber thermoplastic composites. In: Rowell R M, Young R A, Rowell J K (eds.) *Paper and Composites from Agro-Based Resources*. CRC Press, Boca Raton, FL, pp. 377–401

Takase S, Shiraishi N 1989 Studies on composites from wood and polypropylenes. II. *J. Appl. Polym. Sci.* **37**, 645–59

Tjong S C, Xu Y, Meng Y Z 1999 Composites based on maleated polypropylene and methyl cellulosic fiber: Mechanical and thermal properties. *J. Appl. Polym. Sci.* **72**, 1647–53

Törmälä P, Pääkkonen E, Luoto K 1985 Injection molding and extrusion of composites of low density polyethylene and plywood grindings. *J. Appl. Polym. Sci.* **30**, 423–7

Wolcott M P, Englund K 1999 A technology review for wood-plastic composites. In: Wolcott M P, Tichy R J, Bender D (eds.) *33rd Int. Particleboard/Composite Mater. Symp. Proc.* Washington State University, Pullman, WA

Woodhams R T, Law S, Balatinecz J J 1993 Intensive mixing of wood fibers with thermoplastics for injection-molded composites. In: Wolcott M P (ed.) *Wood Fiber/Polymer Composites: Fundamental Concepts, Processes and Material Options*. Forest Products Society, Madison, WI

Woodhams R T, Thomas G, Rodgers D K 1984 Wood fibers as reinforcing fillers for polyolefins. *Polym. Eng. Sci.* **24**, 1166–71

Young R A, Denes F, Nielsen L, Tu X 1995 Improvement of biobased fiber-plastic composite properties through cold plasma treatments. In: Caulfield D F, Rowell R M, Youngquist J A (eds.) *Woodfiber-Plastic Composites: Virgin and Recycled Wood Fiber and Polymers for Composites*. Forest Products Society, Madison, WI

Zehner B 2000 Vinyl based cellulose reinforced composite. *US Patent 6011091*

M. P. Wolcott
Washington State University, Pullman,
Washington, USA

X

Xerogels

A xerogel is a dry gel ("xero" being a Greek word root meaning "dry") prepared by evaporating the pore liquid under ambient conditions. The structure of the solid network is drastically altered during the drying process, because large stresses are exerted by capillary pressure (during evaporation) or crystal growth (during freezing). Those stresses and their effects are the subjects of Sect. 1. The structure and properties of the dried gel are discussed in Sect. 2 and applications are briefly reviewed in Sect. 3. A much more detailed discussion is given by Brinker and Scherer (1990). For information on related materials see *Organic Aerogels** and *Inorganic and Inorganic–Organic Aerogels**

1. *Preparation of xerogels*

1.1 *Shrinkage During Evaporation*

The specific energy of the solid/vapor interface, γ_{SV}, is greater than that of the solid/liquid interface, γ_{SL}, or the liquid/vapor interface, γ_{LV}. Therefore, rather than allowing the solid network to be exposed to the atmosphere during evaporation, it is energetically favorable to increase the area of the liquid/vapor interface by creating concave menisci, as indicated in Fig. 1. The curvature of the liquid/vapor surface creates negative capillary pressure (suction), P_c, in the liquid given by

$$P_c = \frac{2\gamma_{LV}}{r_m} \quad (1)$$

where r_m is the radius of the meniscus; for a concave meniscus, r_m is defined as negative. It is the suction in the liquid that causes the solid network to contract during drying. Initially, the soft gel contracts under small pressures; i.e., r_m is very large and the menisci are relatively flat (Fig. 1(a)). As the gel shrinks, it becomes increasingly rigid, so the menisci deepen and P_c increases. As the bulk modulus of the network, K, increases, the radii of the menisci finally become equal to the radii of the pores, and the liquid recedes into the interior of the network (Fig. 1(c)). At that point, the gel stops shrinking, and one can say that a xerogel has been created.

Even the lowest density inorganic gels have average pore sizes smaller than 100 nm (Ruben *et al.* 1995), and xerogels typically have pore radii from 1 nm to 10 nm, so the capillary pressures generated during drying are large. If $\gamma_{LV} = 0.02\,\mathrm{J\,m^{-2}}$ (typical of organic

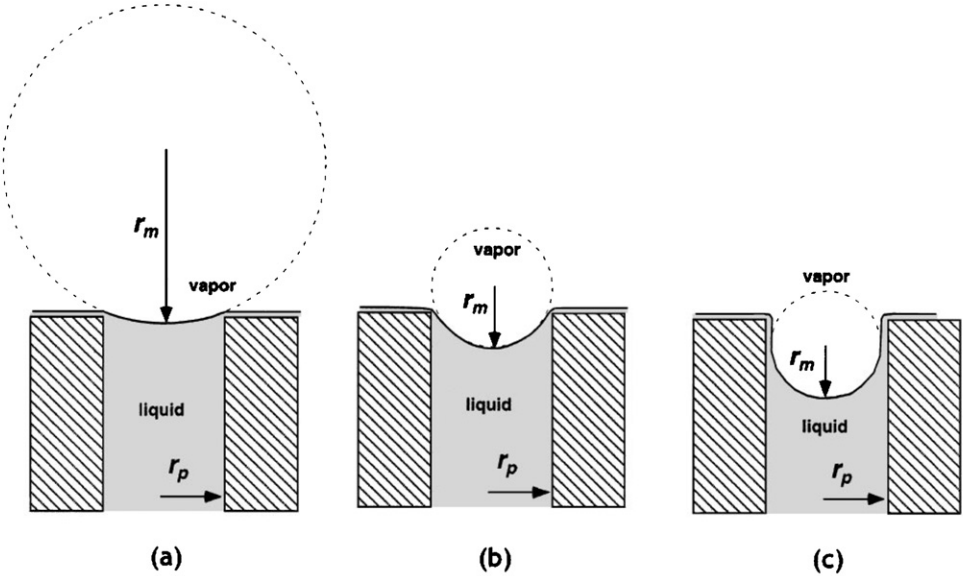


Figure 1
Concave menisci form at surface of drying body. (a) Initially the curvature is small, because the compliant network contracts under small capillary pressure; (b) as the network stiffens, the curvature of the meniscus increases, until (c) it is small enough to penetrate the pore; shrinkage stops when stage (c) is reached.

---

* Cross references marked by an asterisk are included in this volume.
# Cross references marked by a hash can be found by consulting the Encyclopedia of Materials: Science and Technology.

solvents) and $r_m = 10\,\text{nm}$, then Eqn. (1) yields $P_c = 4\,\text{MPa}$; for water, $\gamma_{LV} = 0.072\,\text{J m}^{-2}$, so the same size pore would generate 7.2 MPa of pressure. These large pressures cause 8- to 10-fold volumetric shrinkage during drying. The elastic modulus of the gel network has been found to increase with the volume fraction of solids, ρ, approximately according to

$$K = K_0\left(\frac{\rho}{\rho_0}\right)^m \tag{2}$$

where ρ_0 and K_0 are the initial values of ρ and K (before drying starts) and m is roughly equal to 3.5. If the gel shrinks in volume by a factor of 10, then $\rho/\rho_0 = 10$ and Eqn. (2) indicates that the modulus increases by a factor of almost 10 000. The gel is transformed from an initial consistency similar to that of a gelatin dessert or tofu into a solid almost as hard as window glass.

In most cases, drying shrinkage is irreversible, because strained bonds in the network react with the pore liquid, and the broken bonds result in permanent deformation; moreover new bonds are formed in the network as it shrinks, and these bonds prevent the compressed network from springing back. Treatment of silica gels with chlorosilanes covers the solid surface with alkyl groups that are chemically inert, allowing the drying shrinkage to be completely recovered (Prakash *et al.* 1995). The spring back is so complete that such materials retain the density of the original gel network, and so are called "ambient pressure aerogels."

1.2 *Damage During Evaporation*

If the capillary pressure is uniform throughout the gel, then the network is put under a hydrostatic load and it contracts isotropically without any tendency to warp or crack. Damage results only when *gradients* in pressure are created, because nonuniform suction will cause different regions of the gel to contract at different rates.

When liquid evaporates from the surface of a gel, the movement of liquid toward the surface obeys Darcy's law:

$$J = -\frac{D}{\eta}\nabla P \tag{3}$$

where J is the flux of liquid (i.e., the volume passing through a unit area per unit time, giving units of velocity), η is the viscosity of the liquid, and D is the permeability of the network (units of area). The permeability is proportional to the volume fraction of liquid and the square of the average pore size, r_p:

$$D \propto (1-\rho)r_p^2 \tag{4}$$

Until air invades the gel (Fig. 1(c)), the flux to the surface must be equal to the rate of evaporation,

$$V_E = J_{\text{surface}} = -\frac{D}{\eta}\nabla P\big|_{\text{surface}} \tag{5}$$

where V_E is the volume of liquid evaporating per unit area of gel per unit time. Thus, the pressure gradient within the gel near the surface is

$$\nabla P\big|_{\text{surface}} = -\frac{V_E\eta}{D} \tag{6}$$

This equation explains why damage can be avoided by slow drying: if V_E is small, then large pressure gradients are not created. High drying temperatures are favorable, because they reduce η. Equations (4) and (6) indicate why it is so much more difficult to dry gels than conventional ceramics without cracking: the small pores in gels make the permeability very low (in spite of the small volume fraction of solids), so a modest rate of evaporation creates a substantial pressure gradient in a gel. Details on the phenomenology of drying and the analysis of drying stresses are discussed in Brinker and Scherer (1990, chaps. 7 and 8).

1.3 *Films*

Drying of films begins before gelation, so it is difficult to measure, much less to predict, the modulus and pore size that control drying shrinkage. There is no opportunity to age the gel before drying, so the structure of the network must be controlled by manipulation of the clusters in the sol. For example, the pH and aging time of the sol can be used to make clusters that are more or less robust, and thereby influence the final porosity of the film (Brinker and Scherer 1990, chap. 13). As for a bulk gel, the amount of shrinkage is determined by the capillary pressure and the modulus of the network, but the stresses that cause cracking are not caused by a pressure gradient. For a film, the stress results from the fact that adhesion prevents shrinkage in the plane of the substrate. The resulting stresses are enormous and are independent of the drying rate; nevertheless, cracking only occurs when the thickness of the film exceeds a certain critical value, which is of the order of 1 μm for an inorganic film. The reason for this behavior is that a growing crack releases strain energy stored in a certain volume surrounding the crack, but expends energy in creating the new surface of the crack. In a thin film, the volume relaxed is small compared to the new area created, so it is energetically unfavorable for a crack to extend; i.e., if the film is thin enough it *cannot* crack, regardless of the magnitude of the stress (Hu and Evans 1989). It is possible to build up thick films by applying multiple layers, but it is generally necessary to fire each layer to increase its strength before the next one is deposited.

2. *Structure of Xerogels*

Inorganic gels have pore diameters in the mesopore (2–50 nm) and/or micropore ($\leqslant 2$ nm) range, and surface areas ranging from about 100 to $1000\,\mathrm{m^2\,g^{-1}}$. This means that about half of the atoms in the gel are on the surface, so that substantial changes in the composition of the gel can be produced via surface reactions (such as nitridation by ammonia or dehydration by chlorine vapor). The main structural difference between aerogels and xerogels is their pore size and density; if an aerogel is soaked in liquid and dried, it contracts enormously, but its surface area changes by no more than 10%.

For silica xerogels, the lowest density reported is about $0.2\,\mathrm{g\,cm^{-3}}$ ($\sim$90 vol.% porosity) whereas acid-catalyzed gels dried slowly reach $\sim$50% porosity. Substitution of methyl triethoxysilane (MTEOS) for tetraethoxysilane (TEOS) reduces the connectivity of the silica network, and the increased compliance leads to an abrupt drop in porosity when the MTEOS/TEOS ratio exceeds about 60% (Fahrenholtz and Smith 1992); the resulting materials contain nothing but a small volume of micropores. Such collapse of the pore structure can be obtained in pure silica gels during film formation if the clusters are prepared at low pH so that they are highly compliant.

Control of pore size and shape can be achieved by use of surfactant templates (Beck *et al.* 1992). Hierarchical structures can be made by inducing phase separation in the sol before gelation (Nakanishi 1997): a micrometer-scale pore network is created by spinodal decomposition within a matrix that is micro- or mesoporous.

3. *Applications*

Dense xerogel films are used as coatings to protect against corrosion and abrasion, whereas porous films are used as membranes (in some cases permitting size-selective separation). Sensors can be prepared by incorporating dyes, enzymes (Avnir *et al.* 1994), or other sensitive species within a xerogel applied as a coating on an optical fiber or other substrate. Xerogel films have long been used for passive optical coatings, such as reflective or antireflective films, but there is also tremendous interest in their use for active optical elements (Klein 1994). This application depends on the ability of gels to encapsulate sensitive organic molecules and to preserve their activity.

Xerogel powders have been used for decades for catalyst substrates, owing to their high surface areas. Sol–gel processing has also been used to prepare nuclear fuel pellets and abrasive powders. Sols can be drawn into fibers to prepare xerogel fibers. Monolithic gels are of interest for lens and fiber-optic preforms. These and other applications are reviewed in Brinker and Scherer (1990, chap. 14) and Klein (1988).

Bibliography

Avnir D, Braun S, Lev O, Ottolenghi M 1994 Enzymes and other proteins entrapped in sol–gel materials. *Chem. Mater.* **6**, 1605–14

Beck J S, Vartuli J C, Roth W J, Leonowicz M E, Kresge C T, Schmitt K D, Chu C T-W, Olson D H, Sheppard E W, McCullen S B, Higgins J B, Schlenker J L 1992 A new family of mesoporous molecular sieves prepared with liquid crystal templates. *J. Am. Chem. Soc.* **114** (27), 10834–43

Brinker C J, Scherer G W 1990 *Sol–Gel Science*. Academic Press, Boston, MA

Fahrenholtz W G, Smith D M 1992 Characterization of microporous silica gels prepared from modified silicon alkoxides. In: Hampden-Smith M J, Klemperer W G, Brinker C J (eds.) *Better Ceramics Through Chemistry V*. Materials Research Society, Pittsburgh, PA, pp. 705–10

Hu M S, Evans A G 1989 The cracking and decohesion of thin films on ductile substrates. *Acta Metall.* **37** (3), 917–25

Klein L C (ed.) 1988 *Sol–Gel Technology for Thin Films, Fibers, Preforms, Electronics, and Specialty Shapes*. Noyes, Park Ridge, NJ

Klein L C 1994 *Sol–Gel Optics, Processing and Applications*. Kluwer, Boston, MA

Nakanishi K 1997 Pore structure control of silica gels based on phase separation. *J. Porous Mater.* **4**, 67–112

Prakash S S, Brinker C J, Hurd A J, Rao S M 1995 Silica aerogel films prepared at ambient pressure by using surface derivatization to induce reversible drying shrinkage. *Nature* **374**, 439–43

Ruben G C, Hrubesh L W, Tillotson T M 1995 High resolution transmission electron microscopy nanostructure of condensed-silica aerogels. *J. Non-Cryst. Solids.* **186**, 209–18

G. W. Scherer
Princeton University, New Jersey, USA

This Page Intentionally Left Blank

Z

Zeolites

The word "zeolite" comes from two Greek words meaning "boiling stones." The name was given in 1756 by a Swedish mineralogist, Axel von Cronstedt, to a mineral which released water visibly when heated (the property of intumescence). The mineral later acquired the name stilbite.

Since that date many similar minerals have been discovered and the word zeolite now describes a family of minerals. They are commonly formed as a result of volcanic activity, either by the action of meteoric water in the environs of a volcano, or by the alteration of wind-blown volcanic dusts deposited in saline lakes. The latter mode of formation has created massive deposits on the earth's surface. Some of these deposits are mined commercially as building stone, and for other uses such as soil remediation agents, waste-water treatment, and animal feed supplements.

Historically zeolites were defined as aluminosilicates possessing three-dimensional frameworks of linked silicon–aluminum–oxygen tetrahedra. This restrictive definition has been widened to include *all* crystalline materials characterized by frameworks of linked tetrahedra, each consisting of four oxygen atoms surrounding a cation which can be Al, Si, P, Be, Zn, Mg, Co, B, etc. The frameworks contain channels and cavities in which are found water molecules and cations. The water can be lost reversibly and the cations can usually undergo facile ion exchange. The channels are in the microporous region and are large enough to accommodate guest molecules, which can be organic molecules.

A further extension to the original definition has been to accept some frameworks which are incomplete in so far as some of the apical oxygen atoms to the tetrahedra are occupied by OH or F groups.

Commonly conference reports and scientific journals will consider frameworks containing octahedra (e.g., titanosilicates) and mesoporous substances (e.g., MCM-41; a Mobil company code) alongside the more traditional zeolites because they have related properties. This article will not include these materials.

1. *Classification*

1.1 Zeolites Which Are Natural Minerals

The International Mineralogical Association has a Commission on New Minerals and Mineral Names which has suggested an approach to natural zeolite classification. The zeolites can be grouped into *series* within which various cation compositional ranges can occur. These are listed in Table 1 with the appropriate cations. They are all aluminosilicates with complete frameworks. In addition, zeolite *species* can be recognized. A full list of these can be found in Coombs *et al.* (1998).

1.2 Synthetic Zeolites

The foregoing classification does not include many zeolite phases which have been synthesized in the laboratory. In some cases the phases are identical in structure to a natural analogue but many have no (currently) known natural counterpart. This may change and there is an increasing number of zeolites first recognized as a synthetic species which have subsequently been found in nature. A full list of synthetic zeolites will not be included for reasons of space. Table 2 gives some illustrations of zeolites which can be made with differing atoms in their tetrahedral positions, including some initially identified from laboratory synthesis but which have later been found as natural minerals.

A definitive list of natural and synthetic zeolite structures, with their idealized compositions and full crystallographic data, has been compiled by Meier

Table 1
Zeolite series, their three-letter codes, and compositional variants (species).

Series	Code	Cation forms (species)	Series	Code	Cation forms (species)
Brewsterite	BRE	Sr,Ba[a]	gmelinite	GME	Na,Ca,K
Chabazite	CHA	Ca,Na,K	heulandite	HEU	Ca,Na,K,Sr
Clinoptilolite	HEU	K,Na,Ca	levyne	LEV	Ca,Na
Dachiardite	DAC	Ca,Na	paulingite	PAU	K,Ca
Erionite	ERI	Na,K,Ca	phillipsite	PHI	Na,Ca,K
Faujasite	FAU	Na,Ca,Mg	stilbite	STI	Ca,Na
Ferrierite	FER	Mg,K,Na			

[a] As an example of the link between zeolite series and zeolite species, the names of the species of zeolites in the brewsterite series will be brewsterite–Sr and brewsterite–Ba.

---

* Cross references marked by an asterisk are included in this volume.

# Cross references marked by a hash can be found by consulting the Encyclopedia of Materials: Science and Technology.

Table 2
Synthetic zeolite frameworks containing various tetrahedral atoms.

Framework	Code	Material[a]
Chabazite	CHA	$AlPO_4$-34, CoAPO-44, CoAPO-47, $GaPO_4$, MeAPO-47, MeAPSO-47, SAPO-34, SAPO-47
Faujasite	FAU	SAPO-37, beryllophosphate X, zincophosphate X,
Pahasapite	RHO[b]	berylloarsenate RHO, beryllophosphate RHO, gallosilicate ECR-10

[a] The symbols refer to synthetic phases: A = Al, P = phosphate, S = Si, Me = Mn, Mg. Other numbers and codes are synthesis batch numbers or company codes. [b] RHO is the code for a synthetic zeolite structurally characterized before pahasapite was found in nature.

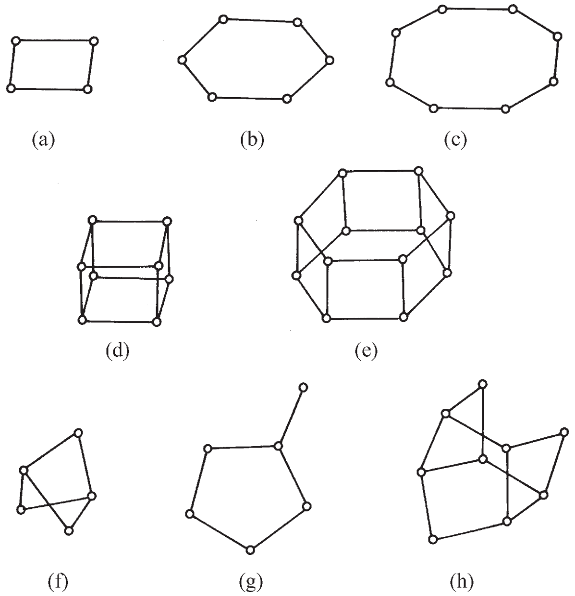


Figure 1
The secondary building units recognized in zeolite frameworks: (a) single four-oxygen ring (S4R), (b) single six ring (S6R), (c) single eight ring (S8R), (d) double four ring (D4R), (e) double six ring (D6R), (f) complex 4–1, (g) complex 5–1, and (h) complex 4–4–1.

et al. (1996). Each framework structure has been allocated a three-letter code by the Structure Commission of the International Zeolite Association. The Structure Commission maintains a website with structural information.

2. *Structural Architecture*

The molecular structures of zeolites are characterized by the unique ways in which their component tetrahedra are linked to each other to create open microporous frameworks. The void volume of some zeolites can be as high as 50% of the total crystal structure. They contain repeating elements of structure which are known as secondary building units (sbu). These are shown in Fig. 1. It is common to simplify the tetrahedral linkages to lines which represent the oxygen atoms, in much the same way as lines are used to represent the tetrahedral carbon bonds in organic molecules. This is illustrated in Fig. 2. This method of representation highlights the ways in which the tetrahedra make rings of oxygen atoms, and the resulting restrictions to the channels and cavities making up the interior void of the structure. These restrictions are of molecular dimensions. They are responsible for the ability of zeolites to act as molecular sieves and perform selective sorption of molecules on the basis of their size to achieve useful separations. Similarly cation-sieving can occur in certain structures.

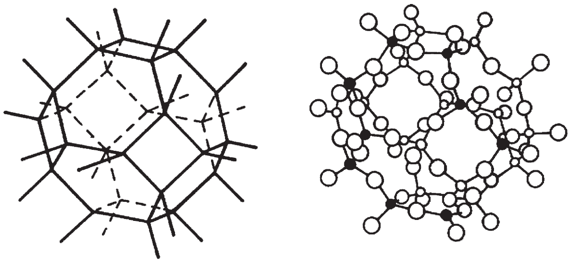


Figure 2
A "ball and stick" representation of the structure of a sodalite unit (β cage) in which the tetrahedra can be indentified (left) with a framework diagram for comparison.

3. *Synthesis*

Earlier methods used to synthesize zeolites tended to mimic their formation in nature by making use of hydrothermal conditions to crystallize out zeolite phases from aqueous gels composed of aluminum and silica sources under highly alkaline conditions. Alteration of Si/Al compositions, amount of water, and differing cation(s) coupled with temperature, autogenous pressure, and time variations resulted in a wide range of products.

Later many zeolites have been synthesized by similar routes but with the inclusion of an organic molecule in the gel as a template, or "void-filler." This approach has generated a number of zeolites with low aluminum contents in their tetrahedral sites (high-silica zeolites).

4. *Industrially Significant Zeolites*

4.1 Synthetic Zeolite A (LTA)

The structure of zeolite A, which has no natural counterpart, is shown in Fig. 3. It is created by the joining together of subunits which are the main structural unit of sodalite—a felspathoid. These units are joined by all their rings of four oxygen atoms into a cubic array like that adopted by the sodium chloride structure. Fig. 3 shows that the consequence of this is a series of cavities—those from the sodalite cage (Fig. 2) itself and those in the central cavities they surround. Entry to the sodalite cage (β-cage) is via a six-oxygen window of 0.23 nm approximate diameter. That to the main cavity (α-cage) is through an eight-oxygen window with a diameter of 0.41 nm. The void volume is 48%. As synthesized, the cavities are occupied by sodium ions and water molecules.

This is the commercial product Molecular Sieve 4A, well-known as an effective laboratory desiccant. Zeolite crystals (about 5 μm) are bound together with a clay (e.g., 10–20% of attapulgite or kaolinite) to form beads or extrudates. They are pre-activated to remove the inherent water prior to use.

The cation present can be changed by cation exchange which moderates the effective eight-oxygen window size. Introduction of the larger potassium cation reduces the working diameter to about 0.3 nm in Molecular Sieve 3A. Replacement of about 40% of the sodium ions by calcium results in a widening to about 0.5 nm, and hence Molecular Sieve 5A. This increase in effective window size arises from the positioning of the calcium cations away from one which constricts the 8-oxygen ring.

(*a*) *Desiccant uses.* Zeolite A is widely used as a desiccant in sealed units for refrigerant gases, in heavy duty oil-filled transformers, double glazing units, vehicle air-brakes, and air conditioning systems.

On a larger scale, zeolite beads are used to dry industrial feedstock gases, such as hydrogen and oxygen. The processes are dynamic with continuous regeneration via temperature- or pressure-swing. Both 3A and 4A are used to dry liquid propane gas (LPG), halocarbons, and natural gas. 3A is preferred for the drying of "cracked" gas (refinery olefin streams), ethylene, propylene, and methanol. 4A dries other alcohols and benzene. It also is preferred for natural gas drying when "peak-shaving" is required. This is the adjustment of the gas calorific value by removing water and carbon dioxide.

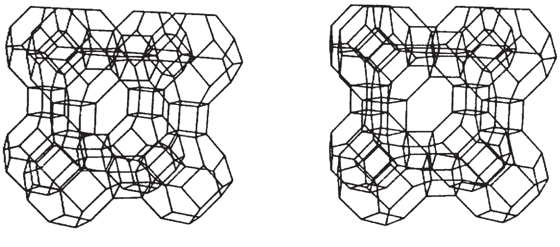


Figure 3
The structure of zeolite A (LTA) as a stereopair.

The basis of the desiccant action described above relies on the scavenging of the water molecule on the basis of its size and polar nature in the presence of other larger/less polar molecules—a molecular sieving.

(*b*) *Gas purification.* The best known of all gas purifications employing a zeolite substrate is in natural gas treatment where, in addition to the peak-shaving mentioned above, 4A "sweetens" the gas by removing hydrogen sulfide. There are many other examples of similar removals of sulfur and nitrogen compounds from gas streams.

(*c*) *Bulk separations.* The window size of 5A is convenient to execute the important separation of *n*-paraffins from *i*-paraffins. Many commercial processes have been developed to carry this out as part of feedstock preparation for the detergent and chemical industries.

Another large-scale sieving separation comes in the use of zeolite A to enrich the gases of the air (oxygen, nitrogen, and inert gases). This can be carried out cryogenically but pressure-swing processes are gaining in use. Oxygen supplies for patients with breathing difficulties and for field hospitals (as in the Gulf War) can be achieved by pressure-swing operations. Other systems provide oxygen for welding and water-oxygenation, and nitrogen for large-scale chemical plant use. Zeolite A has long been an important gas–solid chromatographic support for the analysis of permanent gases.

(*d*) *Liquid phase separations.* Zeolite 5A can be used in small-scale plants for the Union Oil Products' (UOP) Molex process, which is designed for the separation of *n*-paraffins from *i*-paraffins, naphthalenes, and other aromatics.

(*e*) *Cation sieving.* The use which accounts for the highest industrial production of zeolite 4A is as a powder for inclusion in industrial and domestic detergent packs. It acts as a water softener (builder) selectively to remove calcium (and to a lesser extent magnesium). This accounts for nearly 1 million tons per annum. A synthetic gismondine—zeolite MAP—is a modern competitor to A for this purpose.

Zeolite A also has been used for clean-up after the Three Mile Island accident to remove radioactive caesium selectively from contaminated waters in the reactor.

4.2 Synthetic Zeolites X and Y (FAU)

Zeolites X and Y form a series of structures which are isostructural with the rare natural mineral faujasite. X and Y differ only in silicon/aluminum ratio (and hence cation content) with X having Si/Al ~1.0–1.4 and Y Si/Al ~1.5–3.0.

Their structure resembles that of zeolite A but with the β-cages joined via half of their six-rings to create an diamond-like array (Fig. 4). Access to the main cavity is controlled by a 12-oxygen ring with a crystallographic diameter of 0.74 nm. The void volume is 53%.

(*a*) *Industrial use*. The zeolite X is commercially available as Molecular Sieve 13X (sodium form) and 10X (calcium form). Molecular sieve 13X contributes to gas purification, being able to scavenge larger molecules containing sulfur and nitrogen. It has also been used for air-enrichment and in sealed window units.

Ion-exchanged forms (K, Ba, Sr) of both X and Y can separate liquid mixtures and have been developed by UOP to separate *p*-xylene from its isomers (Parex process), ethylbenzene from its isomers (Ebex), and in the Sarex process to separate fructose from other sugars.

The major use of zeolite Y is as a catalyst of great significance to the oil industry. This is covered elsewhere in this encyclopedia.

5. Other Zeolites of Commercial Significance

5.1 Clinoptilolite (HEU) and Mordenite (MOR)

Both these zeolites have silica-rich frameworks based on five-oxygen rings linked together: 5–1 for mordenite and 4–4–1 for clinoptilolite (Fig. 1). The structures have lower void volumes than those of zeolites A, X, and Y. HEU has channels bounded by two eight-ring windows (0.26 × 0.47 nm, 0.33 × 0.46 nm) and one by a 10-ring (0.30 × 0.76 nm). MOR has eight-and 12-ring windows of 0.26 × 0.57 nm and 0.65 × 0.70 nm dimensions, respectively.

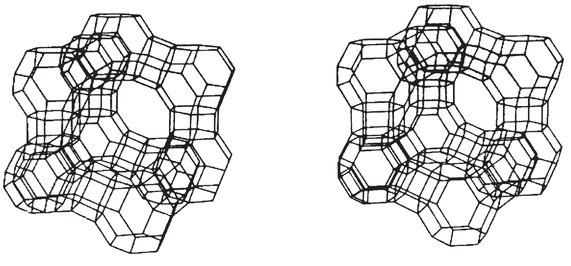


Figure 4
The structure of zeolites X and Y (FAU) as a stereopair.

HEU has not yet been commercially synthesized. It is a very abundant mineral and is mined commercially in many countries. MOR has been used both as a natural mineral and as a synthetic phase.

Their relatively high silicon content enables them to withstand exposure to acid gases and they have been used for acid-gas drying and removal of sulfur and nitrogen oxides. HEU has the useful property of being able to separate nitrogen from natural gas. HEU also has a selectivity for caesium (and strontium). This ion-sieving has been used to clean up after the Chernobyl accident and in the Site Ion-Exchange Plant at British Nuclear Fuel's Sellafield site. It also has selectivity for the ammonium cation which promotes its use in waste-water treatment, soil amendment, and as an animal feed additive.

5.2 Chabazite (CHA)

Although a laboratory synthesis is available for this zeolite, such use that has been made of its properties has been from natural sources. It has found use as a cheap drying agent and as a selective caesuim cation exchanger in the nuclear industry. It has been used in this way in an admixture with zeolite A in the Three Mile Island clean-up and for low-activity waste streams at the Oak Ridge National Laboratory, Tennessee.

5.3 ZSM-5

Zeolite Socony Mobil-5 is a high-silica synthetic zeolite catalyst. It is mentioned here because it is a shape-selective catalyst that acts as a molecular sieve to tailor both reactants and products. This will be described in more detail elsewhere in this encylopedia. The high silica content confers a hydrophobic nature to enable its use in separating organic from aqueous phases, such as in fermentation processes.

6. Other Comments

The nuclear industry has also made use of gas sorption by various zeolites to scavenge and contain tritium, tritiated water, and radioactive krypton from nuclear plant-off gases. Silver-loaded zeolites have been used to scavenge radio-iodine species. Some of these uses have included the deliberate engineering of the pore sizes by deposition of other phases at the pore openings.

7. Conclusion

Most of the studies on the many novel synthetic zeolites have contained an over-emphasis on their potential use as catalysts. The future should see some of them finding wider use as selective sorbents (and cation exchangers) as they offer the prospect of tailoring their molecular sieve properties to meet specific industrial and environmental needs.

Bibliography

Breck D W 1984 *Zeolite Molecular Sieves*. R.A. Krieger, Malabar, FL

Coombs D S, *et al*. 1998 *Recommended Nomenclature for Zeolite Minerals*. Report of the subcommittee on zeolites of the International Mineralogical Association, Commission on New Mineral Names. *Miner. Mag*. **62**, 533–71

Dyer A 1988 *An Introduction to Zeolite Molecular Sieves*. Wiley, Chichester, UK

Karge H G, Weitkamp J (eds.) 1998 *Molecular Sieves, Science and Technology: Synthesis*. Springer, Berlin, Vol. 1

Meier W M, Olson D H, Baerlocher C 1996 Atlas of zeolite structure types. *Zeolites* **17**, 1–230

Structure Commission of the International Zeolite Association, website. http://www.kristall.ethz.ch/IZA-SC/

Szostak R 1984 *Handbook of Molecular Sieves*. Van Nostrand Reinhold, New York

A. Dyer
University of Salford, UK

This Page Intentionally Left Blank

LIST OF CONTRIBUTORS

Contributors are listed in alphabetical order together with their addresses. Titles of articles that they have authored follow in alphabetical order. Where articles are coauthored, this has been indicated by an asterisk preceding the title.

Adcock, T.
Washington State University
Pullman, Washington
USA
**Wood: Nonstructural Panels*

Ahmad, Z.
Quaidi-i-Azam University
Islamabad
USA
Ceramic-modified High-temperature Polymers

Alderliesten, R. C.
Delft University of Technology
The Netherlands
**Fiber Metal Laminates, Fatigue of*

Allison, J. E.
Ford Motor Company
Dearborn, Michigan
USA
**Fatigue of Particle Reinforced Materials*

Ashby, M. F.
University of Cambridge
Cambridge
UK
Composite Materials, Microstructural Design of
**Elastic Structures in Design*
**Materials Selection and Mechanical Design*
Plastic Deformation of Cellular Materials

Avnir, D.
Hebrew University of Jerusalem
Israel
**Reactive Ceramic Nanocomposites with Organic and Bio-organic Phases*

Bailie, J. E.
Cardiff University
UK
**Supported Catalysts*

Baker, R. T. K.
Northeastern University
Boston, Massachusetts
USA
Carbon Nanofibers

Bažant, Z. P.
Northwestern University
Evanston, Illinois
USA
Creep of Concrete

Becher, P. F.
Oak Ridge National Laboratory
Tennessee
USA
Ceramics: Whisker Toughening

Béguelin, Ph.
EPFL
Lausanne,
Switzerland
**Polymer Matrix Composites with Roughly Equiaxed Reinforcements: Microstructure and Mechanical Behavior*

Bergmans, L.
Catholic University of Leuven
Belgium
**Composite Dental Materials: Wear*

Beukers, A.
Delft University
The Netherlands
Polymer Matrix Composites: Applications

Biswas, A.
State University of New York at Buffalo
New York
USA
**Ceramic Nanocomposites with Organic Phases, Optics of*

Blum, J.
Hebrew University of Jerusalem
Israel
**Reactive Ceramic Nanocomposites with Organic and Bio-organic Phases*

Boehner, A. W.
Trus Joist – A Weyerhaeuser Business
Boise, Idaho
USA
Wood: Structural Panel Processes

Boerstoel, H.
Teijin Twaron
Arnhem
The Netherlands
**Processing Rigid Polymers to High Performance Fibers*

Boller, C.
European Aeronautic Defence and Space Company
Muenchen
Germany
Composites for Sensors and Actuators

Bourban, P.-E.
Ecole Polytechnique Fédérale de Lausanne
Switzerland
**Polymer Matrix Composites: Matrices and Processing*

Boyer, R.
Boeing Commercial Airplane Company
Seattle, Washington
USA
Aircraft Materials

Brauns, J.
Latvia University of Agriculture
Jelgava
Latvia
**Elasticity in Wood and Concrete: Hygromechanical Effects*

Broughton, W. R.
National Physical Laboratory
Teddington
UK
**Continuous Parallel Fiber Composites: Deformation and Strength*

Brown, S. F.
University of Nottingham
UK
Creep of Bituminous Aggregates

Buchoff, L. S.
TYCO
Fort Washington, Pennsylvania
USA
Electronic Packaging: Elastomer Conductive Polymers

Bunsell, A. R.
Ecole des Mines de Paris
Evry
France
Composite Materials: Environmental Effects
CVD Monofilaments
High-performance Fibers

Burchell, T. D.
Oak Ridge National Laboratory
Tennessee
USA
Radiation Effects in Carbon–Carbon Composites

Burg, K. J. L.
Clemson University
Clemson, South Carolina
USA
**PES and PEEK*

Buschow, K. H. J.
University of Amsterdam
The Netherlands
Magnets: Bonded Permanent Magnets

Bush, A. L.
Lakewood
Colorado
USA
Construction Materials: Lightweight Aggregates

Calvert, P.
University of Arizona
Tucson, Arizona
USA
Biological Materials: Synthesis, Structure, and Properties
Biomimetic Materials: Properties and Processing

Cameron, N. M.
Owens Corning Science and Technology Center
Granville, Ohio
USA
**Fiberglass*

Casey, H.
Los Alamos National Laboratory
Los Alamos, New Mexico
USA
Sporting Materials: Ski Equipment

Cash, C. G.
Simpson Gumpertz and Heger Inc.
Arlington, Massachusetts
USA
Roofing Materials

Caulfield, D. F.
USDA Forest Service
Madison, Wisconsin
USA
Paper: Effects of Moisture and Temperature

Cervenka, A.
University of Manchester Institute of Science and Technology
UK
Elastic Behavior of Composites

Chawla, K. K.
University of Alabama at Birmingham
Birmingham, Alabama,
USA
Fibrous Reinforcements for Composites: Overview
Glass Fibers

Chawla, N.
Arizona State University
Tempe, Arizona
USA
**Fatigue of Particle Reinforced Materials*

Chermant, J.-L.
LERMAT
Caen
France
**Composites, Microstructure of: Quantitative Description*

Chevalier, J.
INSA
Villeurbanne
France
**Ceramic Matrix Composites with Roughly Equiaxed Reinforcements: Microstructure and Mechanical Behavior*

Claussen, N.
Technische Universität Hamburg–Harburg
Hamburg
Germany
**Reaction Forming*

Clyne, T. W.
University of Cambridge
UK
**Composites: Interfaces*
Metal Matrix Composites: Matrices and Processing

Coffin, D. W.
Institute of Paper Science and Technology
Atlanta, Georgia
USA
**Paper: Creep*
**Paper: Strength and Stiffness*

Cohen, R. E.
Massachusetts Institute of Technology
Cambridge Massachusetts
USA
Block Copolymers as Templates for Functional Materials

Cole, G.
Ford Motor Company
Dearborn, Michigan
USA
Automotive Chassis/Suspension Materials

Conners, T. E.
Forest Products Laboratory
Mississippi
USA
Wood: Ultrastructure

Cooke, T. F.
TRI
Princeton, New Jersey
USA
Fibers: Superabsorbent

Coster, M.
LERMAT
Caen
France
**Composites, Microstructure of: Quantitative Description*

Currey, J. D.
University of York
UK
Bone and Natural Composites: Properties

De Munck, J.
Catholic University of Leuven
Belgium
**Composite Dental Materials: Wear*

Degischer, H. P.
Technische Universität Wien
Austria
Metal Matrix Composites, Recycling of

Donaldson, L. A.
Forest Research
Rotorua
New Zealand
**Wood, Constituents of*

Dyer, A.
University of Salford
UK
Zeolites

Eggeler, G. F.
Ruhr-Universität
Bochum
Germany
Creep and Creep-fatigue of Metal-Matrix Composites

Ellsworth, M. W.
Tyco Electronics
Menlo Park, California
USA
Polymer–Ceramic Nanocomposites: Control of Shrinkage

Evans, P. D.
Australian National University
Canberra, Australian Capital Territory
Australia
Wood Products: Weathering

Fantozzi, G.
INSA
Villeurbanne
France
**Ceramic Matrix Composites with Roughly Equiaxed Reinforcements: Microstructure and Mechanical Behavior*

Fellers, C.
Skogsindustrins Teckniska Forskningsinstitut
Stockholm,
Sweden
**Paper: Creep*
**Paper: Strength and Stiffness*

Fricke, J.
Physikalisches Institut der Universistat
Wurzburg
Germany
**Organic Aerogels*

Friedman, P. A
Ford Motor Company
Dearborn, Michigan
USA
**Automotive Body Materials*

Friedrich, K.
Universität Kaiserslautern
Germany
**Polymer Composites, Friction and Wear of*

Friend, C. S.
State University of New York at Buffalo
New York
USA
**Ceramic Nanocomposites with Organic Phases, Optics of*

Frisch, H. L.
State University of New York at Albany
New York
USA
**Polymers in Constrained Ceramic Environments*

Gin, D. L.
University of California
Berkeley, California
USA
**Organic Analogues to Zeolites and Mesoporous Sieves*

Goehring, T. N.
University of Zurich
Germany
**Composites for Biomedical Applications*

Gooding, R.
Pulp and Paper Research Institute of Canada
Pointe-Claire
Canada
**Papermaking*

Goovaerts, K.
Catholic University of Leuven
Belgium
**Composite Dental Materials: Wear*

Green, D. J.
Pennsylvania State University
University Park, Pennsylvania
USA
Porous Ceramic Processing

Green, D. W.
USDA Forest Service
Madison, Wisconsin
USA
Wood: Strength and Stiffness

Habashi, F.
Laval University
Quebec City, Quebec
Canada
Asbestos

Harmer, M. P.
Lehigh University
Bethlehem, Pennsylvania
USA
**Nanoscale Ceramic Composites*

Hartley, I.
Forintek Canada Corp.
Vancouver, British Columbia
Canada
Wood: Moisture Content, Hygroscopicity, and Sorption

Hearle, J. W. S.
UMIST
Manchester
UK
Textile Fibers: A Comparative Overview

Hedberg, E. L.
Rice University
Houston, Texas
USA
**Bone Augmentation and Repair*

Hedrick, J. L.
IBM Almaden Research Center
San Jose, California
USA
**Block Copolymers as Precursors to Porous Materials*

Heger, F. J.
Simpson Gumpertz & Heger Inc.
Arlington
Massachusetts
USA
**Buildings: Plastics and Composites*

Heppenstall-Butler, M.
University of Manchester Institute of Science and Technology
UK
**Polymer Transcrystallinity in Composites*

Hintz, H. L.
Westvaco Corporation
New York
USA
Paper: Pulping and Bleaching

Houston, D. Q.
Ford Motor Company
Dearborn, Michigan
USA
**Automotive Body Materials*

Huang, L.
State University of New York at Albany
New York
USA
**Polymers in Constrained Ceramic Environments*

Humphrey, P. E.
Oregon State University
Corvallis, Oregon
USA
Wood: Thermal Properties

Hunston, D.
National Institute of Standards and Technology
Gaithersburg, Maryland
USA
**Continuous Parallel Fiber Composites: Fracture*

Hunt, W. H.
Aluminum Consultants Group Inc.
Murrysville, Pennsylvania
USA
Metal Matrix Composites: Applications

Hüsing, N.
Technische Universität Wien
Vienna
Austria
**Inorganic and Inorganic–Organic Aerogels*

Hutchings, G. J.
Cardiff University
UK
**Supported Catalysts*

Iyengar, S.
Saint-Gobain/Higgins Grinding Technology Center
Worcester, Massachusetts
USA
**Conventional and Super Abrasive Materials*

Jin, C.
Texas Instruments
Dallas, Texas
USA
Porous Silica Xerogel

Jones, A. R.
University of Liverpool
UK
Mechanical Alloying

Jones, F. R.
University of Sheffield
UK
**Composites: Interfaces*

Kamke, F. A.
Virginia Polytechnic Institute and State University
Blacksburg, Virginia
USA
Wood: Nonstructural Panel Processes

Karacabeyli, E.
Forintek Canada Corp.
Vancouver, British Columbia
Canada
Wood: Creep and Creep Rupture

Katz, J. D.
Los Alamos National Laboratory
Los Alamos, New Mexico
USA
Whiskers

Kausch, H. H.
EPFL
Lausanne
Switzerland
**Polymer Matrix Composites with Roughly Equiaxed Reinforcements: Microstructure and Mechanical Behavior*

Kikutani, T.
Tokyo Institute of Technology
Japan
Polymer Fibers: Formation and Structure

Klein, L. C.
Rutgers University
Piscataway, New Jersey
USA
**Polymer–Ceramic Nanocomposites: Polymer Overview*

Köhler, L.
University of Freiburg
Germany
Natural Cellulose Fibers: Properties

Komanduri, R.
Mechanical & Aerospace Engineering
Oklahoma State University, Stillwater
Oklahoma
USA
**Conventional and Super Abrasive Materials*

Krause, A. R.
Ford Motor Company
Dearborn, Michigan
USA
**Automotive Body Materials*

Kraynik, A. M.
Sandia National Laboratories
Albuquerque, New Mexico
USA
**Elastic Behavior of Cellular Solids*
**Foams, Microrheology of*

Kumudinie, C.
University of Cincinnati
Ohio
USA
Polymer–Ceramic Nanocomposites: Interfacial Bonding Agents

Künzi, H. U.
EPFL
Lausanne
Switzerland
Metallic Filaments

Lackey, W. J.
Georgia Institute of Technology
Atlanta, Georgia
USA
Carbon–Carbon Composites

Lagace, P. A.
Massachusetts Institute of Technology
Cambridge, Massachusetts
USA
**Designing with Composites*

Lambrechts, P.
Catholic University of Leuven
Belgium
**Composite Dental Materials: Wear*

Lebeau, B.
Université Pierre et Marie Curie
Paris
France
**Optical Properties of Functional Hybrid Organic–Inorganic Nanocomposites*

Lee, P. J.
The University of Wisconsin
Madison, WI
USA
Superconducting Wires and Cables: Materials and Processing

Lemstra, P. J.
Dutch Polymer Institute/Eindhoven University of Technology
Eindhoven
The Netherlands
**Processing Flexible Polymers to High Performance Fibers*

Leung, C. K. Y.
Hong Kong University of Science and Technology
Kowloon
Hong Kong
Concrete as a Building Material

Lev, O.
Hebrew University of Jerusalem
Israel
**Reactive Ceramic Nanocomposites with Organic and Bio-organic Phases*

Lloyd, D.
Alcan International Ltd.
Kingston, Ontario
Canada
Metal Matrix Composites with Roughly Equiaxed Reinforcements: Microstructure and Mechanical Behavior

Lyons, A. M.
Lucent Technologies
Murray Hill, New Jersey
USA
Electronic Packaging: Conductive Adhesives

Maine, E.
Engineering Design Centre
Cambridge
UK
**Materials Selection and Mechanical Design*

Majumdar, B. S.
New Mexico Tech
Socorro New Mexico
USA
**Continuous Parallel Fiber Composites: Fracture*

Manas-Zloczower, I.
Case Western Reserve University
Cleveland, Ohio
USA
**Polymer Melt Mixing: Agglomerate Dispersion*

Manocha, L. M.
Sardar Patel University
Gujarat
India
Carbon Fibers

Månson, J.-A. E.
Ecole Polytechnique Fédérale de Lausanne
Switzerland
**Polymer Matrix Composites: Matrices and Processing*

Mari, D.
Swiss Federal Institute of Technology
Lausanne
Switzerland
Cermets and Hardmetals

Mark, J. H.
University of Cincinnati
Ohio
USA
Elastomers, Ceramic-modified

Maschmeyer, Th.
Delft University of Technology
The Netherlands and
Imperial College of Science
Technology and Medicine
London
UK
**Mesoporous Molecular Sieves*

Matthews, F. L.
Imperial College of Science
Technology and Medicine, London
UK
Composites, Joining of

Mayer, A. B. R.
Polytechnic University Brooklyn
New York
USA
Polymer Nanocomposites with Metal Dispersions

Mayer, J.
Swiss Federal Institute of Technology
Schlieren
Switzerland
**Composites for Biomedical Applications*

McCartney, L. N.
National Physical Laboratory
Teddington
UK
**Continuous Parallel Fiber Composites: Deformation and Strength*
Laminates: Physical and Mechanical Behavior

McDonald, A. G.
Forest Research
Rotorua
New Zealand
**Wood, Constituents of*

Medalia, A. I.
Newton
Massachusetts
USA
Elastomers, Reinforcement of

Mikos, A. G.
Rice University
Houston, Texas
USA
**Bone Augmentation and Repair*

Miller, S. A.
University of California
Berkeley, California
USA
**Organic Analogues to Zeolites and Mesoporous Sieves*

Mitzi, D. B.
IBM T.J. Watson Research Center
Yorktown Heights, New York
USA
Organic–Inorganic Composite Crystals

Moslemi, A. A.
University of Idaho
Moscow
USA
Wood Composites: Mineral-bonded

Naaman, A. E.
University of Michigan
Ann Arbor, Michigan
USA
Reinforced Concrete

Naslain, R. R.
Université Bordeaux 1
Pessac
France
Ceramic Matrix Composites: Matrices and Processing

Nayler, P.
Nayler Chemicals Ltd.
Wigan
UK
Bitumens: Modified

Neilsen, M. K.
Sandia National Laboratories
Albuquerque, New Mexico
USA
**Elastic Behavior of Cellular Solids*

Neubrand, A.
Technische Universität Darmstadt
Germany
Functionally Graded Materials

Nissan, A. H.[†]
Scarsdale
New York
USA
Paper: Effects of Moisture and Temperature

Northolt, M. G.
AKZO Nobel
Arnhem
The Netherlands
**Processing Rigid Polymers to High Performance Fibers*

O'Halloran, M. R.
Western Wood Products Association
Portland, Oregon
USA
Wood: Structural Panels

Okamura, K.
Osaka Prefecture University
Osaka
Japan
Ceramic Fibers from Polymer Precursors

O'Leary, S.
Cardiff University
UK
**Supported Catalysts*

Oppermann, W.
Universität Stuttgart
Germany
Textile Fibers: Mechanical Properties

Pekala, R. W.
ENTEK International LLC
Lebanon, Oregon
USA
**Organic Aerogels*

Perkins, R.
Syracuse University
New York
USA
Paper: Structure

Phulé, P. P.
University of Pittsburgh
Pennsylvania
USA
**Ceramics and Glasses, Sol–Gel Synthesis of*

Picken, S. J.
Delft University of Technology
The Netherlands
**Processing Rigid Polymers to High Performance Fibers*

Pikulik, I. I.
Pulp and Paper Research Institute of Canada
Pointe-Claire
Canada
**Papermaking*

Plummer, C. J. G.
Ecole Polytechnique Fédérale de Lausanne
Switzerland
**Polymer Matrix Composites: Matrices and Processing*

Poirier, N. A.
Pulp and Paper Research Institute of Canada
Pointe-Claire
Canada
**Papermaking*

Prasad, P. N.
State University of New York at Buffalo
New York
USA
**Ceramic Nanocomposites with Organic Phases, Optics of*

Pratt, G. C.
Isle of Harris
UK
Bearing Materials: Plain Bearings

Premachandra, J. K.
University of Cincinnati
Ohio
USA
Polymer–Nonsilica Ceramic Nanocomposites

Prewo, K. M.
United Technologies Research Center
East Hartford, Connecticut
USA
**Ceramic Matrix Composites: Applications*

Pu, Z.
Cincinnati Financial Corporation
Ohio
USA
Membrane-based Polymer–Ceramic Nanocomposites

Pukánszky, B.
Technical University of Budapest
Hungary
Mineral-filled Polymers

Raimondi, M. E.
Imperial College of Science
Technology and Medicine
London
UK
**Mesoporous Molecular Sieves*

Ramamurty, U.
Indian Institute of Science
Bangalore
India
Mechanical Testing Methods of Fibers and Composites

Rånby, B.†
Royal Institute of Technology
Stockholm
Sweden
Natural Cellulose Fibers and Membranes: Biosynthesis

Rapp, C. F.
Owens Corning Science and Technology Center
Granville, Ohio
USA
**Fiberglass*

Rastogi, S.
Dutch Polymer Institute/Eindhoven University of Technology
Eindhoven
The Netherlands
**Processing Flexible Polymers to High Performance Fibers*

Rawski, D. P.
Port Edwards
Wisconsin
USA
Pulp and Paper: Nonfibrous Components

Reinelt, D. A.
Southern Methodist University
Dallas, Texas
USA
**Foams, Microrheology of*

Reinicke, R.
Universität Kaiserslautern
Germany
**Polymer Composites, Friction and Wear of*

Rho, J.-Y.
University of Memphis
Tennessee
USA
Hard Tissues, Mechanical Properties of

Riman, R. E.
Rutgers University
Piscataway, New Jersey
USA
Powders, Solution Synthesis of

Rocens, K.
Riga Technical University
Latvia
**Elasticity in Wood and Concrete: Hygromechanical Effects*

Rodgers, M. B.
The Goodyear Tire & Rubber Company
Akron, Ohio
USA
Rubber Tires

Rösler, J.
Technische Universität Braunschweig
Germany
Creep of Particle Reinforced Materials

Russell, T. P.
University of Massachusetts
Amherst
USA
**Block Copolymers as Precursors to Porous Materials*

Sanchez, C.
Université Pierre et Marie Curie
Paris
France
**Optical Properties of Functional Hybrid Organic–Inorganic Nanocomposites*

Sandhage, K. H.
Ohio State University
Columbus, Ohio
USA
**Reaction Forming*

Sasaki, H.
Akita Prefectural University
Akita
Japan
Lumber: Laminated Veneer

Scherer, G. W.
Princeton University
New Jersey
USA
Xerogels

Schmidt, H.
Universitt des Saarlandes
Saarbrucken
Germany
**Polymer-based Nanocomposites by Sol–Gel Routes, Applications of*

Schneider, M. H.
University of New Brunswick Fredericton
NB
Canada
Polymer Modified Wood

Schubert, U.
Technische Universitt Wien
Vienna
Austria
**Inorganic and Inorganic–Organic Aerogels*

Sebok, E. B.
Columbian Chemicals Company
Marietta, Georgia
USA
**Carbon Blacks*

Seddon, J. M.
Imperial College of Science
Technology and Medicine, London
UK
**Mesoporous Molecular Sieves*

Shalaby, S. W.
PolyMed Inc.
Anderson, South Carolina
USA
**PES and PEEK*

Sharff, P. A.
Simpson Gumpertz & Heger, Inc.
Arlington, Massachusetts
USA
**Buildings: Plastics and Composites*

Shercliff, H. R.
University of Cambridge
UK
**Elastic Structures in Design*

Sherman, A. M.
Ford Motor Company
Dearborn, Michigan
USA
**Automotive Body Materials*

Sleight, A. W.
Department of Chemistry
Oregon State University
Corvallis, OR
USA
Thermally Contracting Materials

Spearing, S. M.
Massachusetts Institute of Technology
Cambridge, Massachusetts
USA
**Designing with Composites*

Stadtwald-Klenke, M.
Universitt des Saarlandes
Saarbrucken
Germany
**Polymer-based Nanocomposites by Sol–Gel Routes, Applications of*

Steenkamer, D. A.
Ford Motor Company
Dearborn, Michigan
USA
**Automotive Body Materials*

Subramoney, S.
DuPont Company
Wilmington, Delaware, USA
Carbon Nanotubes

Sunderland, P.
Swiss Federal Institute of Technology (EPFL)
Lausanne
Switzerland
Polymer Matrix Composites, Recycling of

Tamon, H.
Department of Chemical Engineering
Kyoto University
Japan
Carbon Aerogels

Taylor, R. L.
Columbian Chemicals Company
Marietta, Georgia
USA
**Carbon Blacks*

Tenne, R.
Department of Materials & Interfaces
Weizmann Institute, Rehovoth
Israel
Inorganic Nanotube Materials

Terentjev, E.
University of Cambridge
UK
Liquid Crystalline Polymers, Dispersed

Thompson, G. S.
Lehigh University
Bethlehem, Pennsylvania
USA
**Nanoscale Ceramic Composites*

Tredway, W. K.
United Technologies Research Center
East Hartford, Connecticut
USA
**Ceramic Matrix Composites: Applications*

Tucker, C. L. III
University of Illinois
Urbana, Illinois
USA
Fiber Orientation in Composite: Influence of Fluid Flow

Ulm, F.-J.
Massachusetts Institute of Technology
Cambridge, Massachusetts
USA
Construction: Cellular Materials

Usuki, A.
Toyota Central R&D Labs
Aichi
Japan
Clay-based Polymer Nanocomposites

Van Meerbeek, B.
Catholic University of Leuven
Belgium
**Composite Dental Materials: Wear*

van Mier, J. G. M.
Delft University of Technology
The Netherlands
Concrete: Failure Mechanics

Viney, C.
Heriot-Watt University
Edinburgh
UK
Natural Protein Fibers

Vlot, A.
Delft University of Technology
The Netherlands
**Fiber Metal Laminates, Fatigue of*

Wang, S.
DuPont Dow Elastomers
Elkton, Maryland
USA
Polymer–Ceramic Nanocomposites: Catalysts

Waterhouse, J. F.
Peachtree City
Georgia
USA
Paper Products: Classification

Wei, Y.
Drexel University
Philadelphia, Pennsylvania
USA
Polymer-modified Ceramics

Wen, J.
Cabot Corporation
Tuscola, Illinois
USA
Polymer–Silica Nanocomposites

Wheeler, E.
North Carolina State University
Raleigh, North Carolina
USA
Wood: Macroscopic Anatomy

Wilcox, W. W.
University of California
Berkeley, California
USA
Wood Products: Decay During Use

Williamson, A.
Cranfield University
UK
Alternative Techniques for Manufacturing Composite Aerospace Designs, Cost of

Wilson, D. M.
3M Corporation
St. Paul, Minnesota
USA
Fiber Strength
Spun (Slurry and Sol–Gel) Ceramic Fibers

Wintermantel, E.
Technical University Munich
Germany
**Composites for Biomedical Applications*

Withers, P. J.
Manchester Materials Science Centre
UK
Composites, Physical Properties of

Wojcik, A. B.
Hybrid Glass Technologies
Princeton, New Jersey
USA
**Polymer–Ceramic Nanocomposites: Polymer Overview*

Wolcott, M. P.
Washington State University
Pullman, Washington
USA
**Wood: Nonstructural Panels*
Wood-Plastic Composites

Wood, T. E.
3M Corporation
St. Paul, Minnesota
USA
**Ceramics and Glasses, Sol–Gel Synthesis of*

Xu, P.
W. L. Gore & Associates Inc.
Elkton, Maryland
USA
Polymer–Ceramic Nanocomposites: Ceramic Phases

Young, J. F.
University of Illinois
Urbana, Illinois
USA
Portland Cements

Young, R. J.
University of Manchester Institute of Science and Technology
UK
**Polymer Transcrystallinity in Composites*

Zweben, C.
Devon
Pennsylvania
USA
Electronic Packaging: Heat Sink Materials